
Release 3.1A John F. Collins, Biocomputing Research Unit.
Copyright (c) 1993-1998 University of Edinburgh, U.K.
Distribution rights by Oxford Molecular Ltd

search on n.a. - n.a. database search. using Smith-Waterman algorithm

```
Run on: Thu Apr 29 02:55:57 1999; MasPar time 500.39 Seconds
1416.800 Million cell updates/sec
Tabular output not generated.
```

```

Title:
Description:
Perfect Score: 299
N.A. Sequence: 1 gctggagagaacacagcaa.....atagctgtgaacacagaa 300
Comp: ccacacctctcttctgtgtt.....tatcgacactctgtgtctt

```

Scoring table: TABLE default
Gap 6

Nmatch STD : Dbase 0: Query 0

Searched: 602357 seqs, 1181590623 bases x 2

Post-processing: Minimum Match 0% Listing first 45 summaries

```

Database:
emb157
1:em_ba 2:em_fun 3:em_htg 4:em_hum1 5:em_hum2 6:em_in
7:em_om 8:em_or 9:em_ov 10:em_pat 11:em_ph 12:em_pl
13:em_ro 14:em_un 15:em_v1
genbank110
16:gb_ba1 17:gb_ba2 18:gb_htg 19:gb_in 20:gb_om 21:gb_
22:gb_pat 23:gb_ph 24:gb_pl1 25:gb_pl2 26:gb_pr1
27:gb_pr2 28:gb_pr3 29:gb_ro 30:gb_st 31:gb_sts 32:gb_
33:gb_un 34:gb_v1

```

Statistics: Mean 9.536: Variance 4.385: scale 2.175

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query		ID	Description	Pred. No.
		Match	Length			
c 1	53	17.7	7218 22	166494	Sequence 14 from paten	5.68e-22
c 2	35	11.7	2215 22	128278	Sequence 5 from patent	1.76e-08
c 3	30	10.0	74371 27	AC005369	Homo sapiens chromosom	4.69e-05
c 4	28	9.4	965 22	AR024229	Sequence 22 from paten	9.58e-04
c 5	27	9.0	215 22	128278	Sequence 5 from patent	4.18e-03
c 6	27	9.0	7332 26	HUMTAN1	Human TAN-1 mRNA (homo	4.18e-03
c 7	27	9.0	183338 18	HS329A5	Human DNA sequence ***	4.18e-03
c 8	27	9.0	216021 27	HUAC004787	Homo sapiens Chromosom	4.18e-03
c 9	26	8.7	965 22	AR024229	Sequence 22 from paten	1.78e-02
c 10	26	8.7	2976 29	MUSENP1A	Mus musculus bone morph	1.78e-02
c 11	26	8.7	8064 29	MNNOPTCHA	M.musculus notch-1 mRN	1.78e-02
c 12	26	8.7	8221 29	RNROTCH	R.rattus mRNA homolog	1.78e-02
c 13	26	8.7	74371 27	AC005369	Homo sapiens chromosom	1.78e-02

ALIGNMENTS

1	166494	7218 bp	DNA	PAT	23-DEC-1997
LOCUS	Sequence 14	from patent	US 5670367.		
DEFINITION	166494				
ACCESSION	166494				
NID	92724471				
KEYWORDS					
SOURCE	Unknown.				
ORIGIN	Unknown.				
REFERENCE	Unclassified.				
AUTHORS	1 (bases 1 to 7218)				
TITLE	Dorner, F., Scheiflinger, F. and Falkner, F. Gunter.				
JOURNAL	Recombinant fowipox virus				
FEATURES	Patent: US 5670367-A 14 23-SEP-1997;				
	Location/Qualifiers				
	1. .7218				
	/organism="unknown"				
BASE COUNT	1944 a	1491 c	1486 g	1929 t	368 others
ORIGIN					

Query Match 17.7%; Score 53; DB 22; Length 7218;
Best Local Similarity 0.3%; Pred. NO. 5.68e-22;
Matches 1; Conservative 173; Mismatches 122. Indels

[illegible]

Cp 120 gcaggtttcccggtgctcctggaagcagctgctgattgttggtccacgcactcacc 61

Db 1306 YY 1361

Cp 60 aaacttacatccaggttcgctgctgctgctcagcagctccctgctgtttctctcc 5

RESULT 2

LOCUS I28278 215 bp DNA PAT 30-OCT-1996

DEFINITION Sequence 5 from patent US 5569830.

ACCESSION I28278

NID 91819054

KEYWORDS

SOURCE Unknown.

ORGANISM Unknown.

REFERENCE 1 (bases 1 to 215)

AUTHORS Bennett,A., Labavitch,J.M., Powell,A. and Stotz,H.

TITLE Plant inhibitors of fungal polygalacturonases and their use to control fungal disease

JOURNAL Patent: US 5569830-A 5 29-OCT-1996;

FEATURES

source

1. .215

Location/Qualifiers

/organism="unknown"

BASE COUNT 15 a 8 c 25 g 26 t 141 others

ORIGIN

Query Match 11.7%; Score 35; DB 22; Length 215;

Best Local Similarity 17.1%; Pred. No. 1.76e-08;

Matches 28; Conservative 64; Mismatches 70; Indels 2; Gaps 2;

Db 52 VNDSGHNKYSSANYNGNNGVGAATHTHTVNSGADSKTVTSYNASGTSSSNGGTD 111

Cp 298 ctgtgcttcacagctatactgacagcttcatcgga-catgctcctngagttcacac 240

Db 112 GNRSGADSYGSSKTAAMTSRNRGKTANNVDSRNMGDASVGSKNYKTHAKNSADGKVG 171

Cp 239 gtgacatctggaatgagctggtgccaactgaggaagcaagcactgtgctcctgtgta 180

Db 172 KNGDRNNRYGTGTSKNSVNNCGGNKRDVSSYANNKCGSSCT 215

Cp 179 ttc-acacatctgtgtgcatggccgggggttcattccacact 137

RESULT 3

LOCUS AC005369 74371 bp DNA PRI 01-AUG-1998

DEFINITION Homo sapiens chromosome 5, BAC clone 119j3 (LBNL H175), complete sequence.

ACCESSION AC005369

KEYWORDS g3367505

SOURCE HTG.

ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;

AUTHORS Primates; Catarrhini; Homnidae; Homo.

1 (bases 1 to 74371)

Kimmerly,W., Bondoc,M., Cheng,J., Connolly,K.S., Gunning,K.M.,

Kadner,K., Miguel,T., Miller,C., Pitluck,S., Pollard,M.,

Rojeski,H., Subramanian,S. and Martin,C.H.

Sequencing of human chromosome 5

Unpublished

2 (bases 1 to 74371)

Ricke,D.O.

Large Scale Sequence Analysis and Annotation with the Sequence

Comparison Analysis (SCAN) System

Unpublished

3 (bases 1 to 74371)

Kimmerly,W., Bondoc,M., Cheng,J., Connolly,K.S., Gunning,K.M.,

Davis,C.A., Kadner,K., Miguel,T., Pitluck,S., Pollard,M.,

Rojeski,H., Subramanian,S. and Martin,C.H.

Direct Submission

Submitted (01-AUG-1998) Human Genome Center, DOE Joint Genome

Institute, Lawrence Berkeley National Laboratory, MS 74-157,

Berkeley, CA 94720, U.S.A.

Sequence submitted by:

DOE Joint Genome Institute.

Location/Qualifiers

1. .74371

/organism="Homo sapiens"

/db_xref="taxon:9606"

/map="5q"

/clone="119j3"

/chromosome="5"

/note="LBNL H175"

893. 1030

/rpt_family="Alu"

2295. .2438

/rpt_family="Alu"

2818. .2859

/note="(GT)21"

/rpt_type=tandem

/rpt_unit=GT

Join(3246. .3410,3721. .3828)

/standard_name="RLF"

/note="65% & 69% protein identity GenPept:U022377"

3431. .3724

/rpt_family="Alu"

3707. .3728

/note="(A)22"

/rpt_type=tandem

/rpt_unit=A

4366. .4661

/rpt_family="Alu"

5327. .5602

/rpt_family="Alu"

6586. .6956

/rpt_family="L1"

6647. .6684

/note="(CA)19"

/rpt_type=tandem

/rpt_unit=CA

7113. .7373

/rpt_family="Alu"

complement(7830. .8185)

/standard_name="possible repeat"

8258. .8503

/rpt_family="Alu"

9070. .9387

/rpt_family="Alu"

complement(9740. .9845)

/rpt_family="MER42"

complement(10440. .11015)

/rpt_family="Alu"

11950. .12250

/rpt_family="Alu"

12057. .12085

/note="(A)29"

/rpt_type=tandem

/rpt_unit=A

12365. .12645

/rpt_family="Alu"

13727. .13750

/note="(AC)12"

/rpt_type=tandem

/rpt_unit=AC

13783. .14024

/rpt_family="L1"

14175. .14470

/rpt_family="Alu"

complement(14906. .15259)

/standard_name="possible repeat"

15300. .15613

/rpt_family="Alu"

16671. .16690

/note="(T)20"

/rpt_type=tandem

[illegible]

DEFINITION	Mus musculus bone morphogenetic protein (Bmp-1) mRNA, complete cds.
ACCESSION	L24755
NID	g439606
KEYWORDS	bone morphogenetic protein-1; metalloproteinase.
SOURCE	Mus musculus (strain C57BL/6) embryo cDNA to mRNA.
ORGANISM	Mus musculus
	Eukaryotes; mitochondrial eukaryotes; Metazoa; Chordata; Vertebrata; Eutheria; Rodentia; Sciurognathi; Myomorpha; Muridae; Murinae; Mus.
REFERENCE	1 (bases 1 to 2976)
AUTHORS	Fukagawa,M., Noboru,S., Hogan,B.L.M. and Jones,C.M.
TITLE	Embryonic expression of mouse bone morphogenetic protein-1 (BMP-1) which is related to the Drosophila dorsoventral gene tolloid and encodes a putative astain metalloproteinase
JOURNAL	Dev. Biol. 163, 175-183 (1994)
MEDLINE	94229342
COMMENT	On Jan 6, 1994 this sequence version replaced gi:437434.
FEATURES	Location/Qualifiers
source	1..2976

gene

11

CDS

666 a 824 c 874 g 612 t

count

...
Note: remainder of annotations omitted.

Query Match 8.4%; Score 25; DB 22; Length 565;
Best Local Similarity 28.2%; Pred. No. 7.37e-02;
Matches 31; Conservative 40; Mismatches 37; Indels 2; Gaps 2;
Db 212 YVATCMTGCWDDYCCBGGGTGYRBCCTGYGTGCGARRRYRRYNNYTCBMGNTGYT 271

PN	EP-571911-A.
PD	01-DEC-1993.
PF	24-MAY-1993; 108325.
PA	(BECT) BECTON DICKINSON CO.
PI	Shank DD, Spears PA;
DR	WPI; 93-378844/48.
PT	New oligo:nucleotide probes specific for Mycobacteria - used for detection and amplification of Mycobacteria nucleic acid in samples
PS	Claim 3; Page 14; 23pp; English.
CC	Oligonucleotide probe MK14-A consists of nucleotides 5-95 of MK14 (Q51735). It hybridized to all spp. of mycobacteria tested, but cross reacted to a few non-mycobacterial spp. The probe may be useful as an initial screen for mycobacterial infection.
CC	See also Q51735-45 and Q51747-59.
SQ	Sequence 91 BP; 5 A; 17 C; 15 G; 4 T;
Db	Query Match 12.4%; Score 37; DB 9; Length 91; Best Local Similarity 2.0%; Pred. No. 8.71e-10; Matches 1; Conservative 43; Mismatches 7; Indels 0; Gap 0
Qy	10 gssvhsyyvvhhshhsvhvhhvvhvhsvvvhhvhhvhhvhhvhyhvyvsv 60 : :: : :: : :: : :: : :: : :: : :: : :: : :: : :: : :: : :: : :: : 61 ggtgagtcgtggcgaccacaataatgcagtcttccaggatacacccgg 111
RESULT	4
ID	Q51746 standard; cDNA; 91 BP.
AC	Q51746;
DE	31-MAY-1994 (first entry)
DE	Oligonucleotide probe MK14-A
KW	Oligonucleotide; DNA probe; mycobacteria; disease diagnosis;
KW	ss.
OS	Synthetic.
PN	EP-571911-A.
PD	01-DEC-1993.
PF	24-MAY-1993; 108325.
PA	(BECT) BECTON DICKINSON CO.
PI	Shank DD, Spears PA;
DR	WPI; 93-378844/48.
PT	New oligo:nucleotide probes specific for Mycobacteria - used for detection and amplification of Mycobacteria nucleic acid in samples
PS	Claim 3; Page 14; 23pp; English.
CC	Oligonucleotide probe MK14-A consists of nucleotides 5-95 of MK14 (Q51735). It hybridized to all spp. of mycobacteria tested, but cross reacted to a few non-mycobacterial spp. The probe may be useful as an initial screen for mycobacterial infection.
CC	See also Q51735-45 and Q51747-59.
SQ	Sequence 91 BP; 5 A; 17 C; 15 G; 4 T;
Db	Query Match 12.4%; Score 37; DB 9; Length 91; Best Local Similarity 2.0%; Pred. No. 8.71e-10; Matches 1; Conservative 43; Mismatches 7; Indels 0; Gap 0
Cp	10 gssvhsyyvvhhshhsvhvhhvvhvhsvvvhhvhhvhhvhhvhyhvyvsv 60 : :: : :: : :: : :: : :: : :: : :: : :: : :: : :: : :: : :: : :: : 98 ggaagcatcgtgcattgtttggtcccacgcactcaccacaattacacca 48
RESULT	5
ID	Q70469 standard; DNA; 114 BP.
AC	Q70469;
DE	07-APR-1995 (first entry)
DE	Genetic DNA sequence to generate a random TSAR peptide library.
KW	TSAR; totally synthetic affinity reagent; synthetic; binding domain effector domain; concatenated heterofunctional protein; linker; direct; rapid; detection; screening; treatment; generic; ss.
OS	Synthetic.
FH	key Location/Qualifiers
FT	misc_feature 55..60

```

PR 01-FEB-1993; US-013416.
PR 30-DEC-1993; US-176500.
PR 31-JAN-1994; US-189331.
PR (UYNC-) UNIV NORTH CAROLINA.
PI Fowlkes DM, Kay BK;
DR WPI: 94-279739/34.
DR P-PSDB; R65154.
PT Identifying proteins or peptide(s) which bind a ligand - by
PT screening a recombinant vector library expressing fusion proteins
PT comprising a binding domain and an effector domain
PS Disclosure; Page 35; 25pp; English.
CC Q70468 is a generic DNA sequence used to generate random TSAR (Totally
CC Synthetic Affinity Reagents) peptides.This generic formula can also be
CC represented as follows: X(NNB)Yll(TGC)(NNB)6z(NNB)7(TGC)(NNB)l0Y.X
CC and Y are flanking restriction sites (X is not the same as Y) that are
CC not specified further. Other generic sequences are shown in Q70466-68.
CC Other specific peptides generated by these generic sequences are shown in
CC R65151-54. TSARs are concatenated heterofunctional proteins or peptides
CC comprising at least two functional regions - a binding domain with
CC affinity for a ligand and a second effector peptide portion that is
CC chemically or biologically active.They may further comprise a linker
CC peptide between the 2 domains.The oligonucleotides are also designed so
CC that the expressed peptide contains 2 or 4 cysteine residues positioned
CC in, or flanking, the unpredicted or variant residues. These residues
CC confer some degree of conformational rigidity to the peptides. The TSARs
CC or compons. comprising a TSAR binding domain can be used in vivo to
CC deliver a chemically or biologically active moiety, eg. metal ion,
CC radioisotope, peptide, toxin or enzyme, to the specific target or on the
CC cell. They can also replace the function of macromolecules, eg.
CC monoclonal or polyclonal antibodies and therefore circumvent the need
CC for complex methods of hybridoma formation or in vivo antibody
CC production. The TSARs are easily characterised and have designed activity
CC allowing direct and rapid detection in a screening process.
CC Sequence 114 BP; 0 A; 2 C; 2 G; 2 T;
SQ
Query Match 10.7%; Score 32; DB 12; Length 114;
Best Local Similarity 4.5%; Pred. No. 9.60e-07;
Matches 5; Conservative 31; Mismatches 76; Indels 0; Gaps 0;
Db 3 bnnbnbnbnbnbnbnbnbnbnbnbnbnbtgcnbnbnbnbnbnbnbnbnbnbnbnbnbnbnbn 62
Cp 200 cacttgtagcttcgtgtattcaccatctgttggcatgcgccgggtttcattcca 141
Dd 63 bnnbnbnbnbnbnbnbnbnbnbnbnbnbnbnbnbnbnbnbnbnbnbnbnbnbnbnbnbn 114
Cp 140 cacttcattcacattgactgcagggtttcccggtgatcctcctggaagcatic 89
RESULT 7
ID Q70465 standard; DNA; 114 BP.
AC AC Q70465;
DE OS-APR-1995 (first entry)
DE Generic DNA sequence to generate a random TSAR petide library.
KW TSAR; totally synthetic affinity reagent; synthetic; binding domain;
KW effector domain; concatenated heterofunctional protein; linker;
KW direct; rapid; detection; screening; treatment; generic; ss.
OS Synthetic.
FH Key Location/Qualifiers
FT misc_feature 55..60
FT FT /tag= a
FT FT /note= "this sequence represents 'Z'; 2 can be a
FT FT sequence of 6, 9 or 12 nucleotides (see
FT FT comments)"
PN WO9418318-A.
PD 18-AUG-1994.
PF 01-FEB-1994. U00977.
PR 01-FEB-1993; US-013416.
PR 30-DEC-1993; US-176500.
PR 31-JAN-1994; US-189331.
PA (UYNC-) UNIV NORTH CAROLINA.
PI Fowlkes DM, Kay BK;
DR WPI: 94-279739/34.
DR P-PSDB; R65150 and R65151.

```

PT Identifying proteins or peptide(s) which bind a ligand - by
 PT screening a recombinant vector library expressing fusion proteins
 PS comprising a binding domain and an effector domain
 PS Disclosure, Page 35; 255pp; English.
 CC Q70465 is a generic DNA sequence used to generate random TSAR (Totally
 CC Synthetic Affinity Reagents) peptides. This generic formula can also be
 CC represented as follows: X(NNB)6(TGC)(NNB)12(NNB)14(TGC)(NNB)3Y. X
 CC and Y are flanking restriction sites (X is not the same as Y) that are
 CC not specified further. Other generic sequences are shown in Q70466-68.
 CC Other specific peptides generated by these generic sequences are shown in
 CC R65151-54. TSARs are concatenated heterofunctional proteins or peptides,
 CC comprising at least two functional regions - a binding domain with
 CC affinity for a ligand and a second effector peptide portion that is
 CC chemically or biologically active. They may further comprise a linker
 CC peptide between the 2 domains. The oligonucleotides are also designed so
 CC that the expressed peptide contains 2 or 4 cysteine residues positioned
 CC in, or flanking, the unpredicted or variant residues. These residues
 CC confer some degree of conformational rigidity to the peptides. The TSARs
 CC or compsns. comprising a TSAR binding domain can be used in vivo to
 CC deliver a chemically or biologically active moiety, eg. metal ion,
 CC radioisotope, peptide, toxin or enzyme, to the specific target or on the
 CC cell. They can also replace the function of macromolecules, eg.
 CC monoclonal or polyclonal antibodies and therefore circumvent the need
 CC for complex methods of hybridoma formation or in vivo antibody
 CC production. The TSARs are easily characterised and have designed
 CC activity allowing direct and rapid detection in a screening process.
 SQ Sequence 114 BP; 0 A; 2 C; 2 G; 2 T;

Query Match 10.0%; Score 30; DB 12; Length 114;
 Best Local Similarity 3.6%; Pred. No. 1.46e-05;
 Matches 4; Conservative 31; Mismatches 77; Indels 0; Gaps 0;
 Db 3 bnnbnnbnnbnnbgtcnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnnnnn 62
 Cp 188 ccgtgtgtattcacacatctgtgtgcatgcccgggtttcattccacactcattca 129
 Db 63 bnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnb 114
 Cp 128 tcttgactgcagggtttcccggtgtatctcctggaaagcatctgcattgttg 77

RESULT 8
 ID Q70467 standard; DNA; 114 BP.
 AC Q70467.
 DT 05-APR-1995 (first entry)
 DE Generic DNA sequence to generate a random TSAR peptide library.
 KW TSAR; totally synthetic affinity reagent; synthetic; binding domain;
 KW effector domain; concatenated heterofunctional protein; linker;
 KW direct; rapid; detection; screening; treatment; generic; ss.
 OS Synthetic.
 FH Key Location/Qualifiers
 FT misc_feature 55..60
 FT /*tag= a
 FT /note= "this sequence represents 'Z'; Z can be a
 FT sequence of 6, 9 or 12 nucleotides (see
 FT comments)."
 PN WO9418318-A.
 PD 18-AUG-1994.
 PF 01-FEB-1994; U00977.
 PR 01-FEB-1993; US-013416.
 PR 30-DEC-1993; US-176500.
 PR 31-JAN-1994; US-189331.
 PA (UINC-) UNIV NORTH CAROLINA.
 PI Fowlkes DM, Kay BK;
 DR WPI; 94-279739/34.
 DR P-PSDB; R65153.
 PT Identifying proteins or peptide(s) which bind a ligand - by
 PT screening a recombinant vector library expressing fusion proteins
 PS comprising a binding domain and an effector domain
 PS Disclosure, Page 35; 255pp; English.
 CC Q70467 is a generic DNA sequence used to generate random TSAR (Totally
 CC Synthetic Affinity Reagents) peptides. This generic formula can also be
 CC represented as follows: X(NNB)16(TGC)(NNB)12(NNB)16(TGC)(NNB)1Y. X

CC and Y are flanking restriction sites (X is not the same as Y) that are
 CC not specified further. Other generic sequences are shown in Q70466-68.
 CC Other specific peptides generated by these generic sequences are shown in
 CC R65151-54. TSARs are concatenated heterofunctional proteins or peptides,
 CC comprising at least two functional regions - a binding domain with
 CC affinity for a ligand and a second effector peptide portion that is
 CC chemically or biologically active. They may further comprise a linker
 CC peptide between the 2 domains. The oligonucleotides are also designed so
 CC that the expressed peptide contains 2 or 4 cysteine residues positioned
 CC in, or flanking, the unpredicted or variant residues. These residues
 CC confer some degree of conformational rigidity to the peptides. The TSARs
 CC or compsns. comprising a TSAR binding domain can be used in vivo to
 CC deliver a chemically or biologically active moiety, eg. metal ion,
 CC radioisotope, peptide, toxin or enzyme, to the specific target or on the
 CC cell. They can also replace the function of macromolecules, eg.
 CC monoclonal or polyclonal antibodies and therefore circumvent the need for
 CC complex methods of hybridoma formation or in vivo antibody production.
 CC The TSARs are easily characterised and have designed activity allowing
 CC direct and rapid detection in a screening process.
 SQ Sequence 114 BP; 0 A; 2 C; 2 G; 2 T;

Query Match 10.0%; Score 30; DB 12; Length 114;
 Best Local Similarity 3.6%; Pred. No. 1.46e-05;
 Matches 4; Conservative 31; Mismatches 77; Indels 0; Gaps 0;
 Db 3 bnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnnnnn 62
 Cp 194 tagcttcggtgtattcacacatctgtgtgcatgcccgggtttcattccacactca 135
 Db 63 bnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnb 114
 Cp 134 ttcacattgactgcagggtttcccggtgtatctcctggaaagcatctgcatt 83

RESULT 9
 ID Q70470 standard; DNA; 114 BP.
 AC Q70470.
 DT 10-APR-1995 (first entry)
 DE Generic DNA sequence to generate a random TSAR peptide library.
 KW TSAR; totally synthetic affinity reagent; synthetic; binding domain;
 KW effector domain; concatenated heterofunctional protein; linker;
 KW direct; rapid; detection; screening; treatment; generic; ss.
 OS Synthetic.
 FH Key Location/Qualifiers
 FT misc_feature 55..60
 FT /*tag= a
 FT /note= "encoded by Z (see comments)"
 PN WO9418318-A.
 PD 18-AUG-1994.
 PF 01-FEB-1994; U00977.
 PR 01-FEB-1993; US-013416.
 PR 30-DEC-1993; US-176500.
 PR 31-JAN-1994; US-189331.
 PA (UINC-) UNIV NORTH CAROLINA.
 PI Fowlkes DM, Kay BK;
 DR WPI; 94-279739/34.
 DR P-PSDB; R58378.
 PT Identifying proteins or peptide(s) which bind a ligand - by
 PT screening a recombinant vector library expressing fusion proteins
 PS comprising a binding domain and an effector domain
 PS Disclosure, Page 36; 255pp; English.
 CC Q70470 is a generic DNA sequence used to generate random TSAR (Totally
 CC Synthetic Affinity Reagents) peptides. This generic formula can also be
 CC represented as follows: X(NNB)4(CAC)(NNB)4(CAC)(NNB)8Z(NNB)6(CAC)(NNB)8
 CC -(CAC)2(NNB)Y. X and Y are flanking restriction sites (X is not the same
 CC as Y) that are not specified further. The peptides generated by this and
 CC other generic sequences (Q70471-73) have invariant histidine residues
 CC incorporated into variant sequences. TSARs are concatenated
 CC heterofunctional proteins or peptides, comprising at least two functional
 CC regions - a binding domain with affinity for a ligand and a second
 CC effector peptide portion that is chemically or biologically active. They
 CC may further comprise a linker peptide between the 2 domains. The TSARs
 CC or compsns. comprising a TSAR binding domain can be used in vivo to

CC deliver a chemically or biologically active moiety, eg. metal ion, radioisotope, peptide, toxin or enzyme, to the specific target or on the cell. They can also replace the function of macromolecules, eg. monoclonal or polyclonal antibodies and therefore circumvent the need for complex methods of hybridoma formation or in vivo antibody production. The TSARs are easily characterised and have designed CC activity allowing direct and rapid detection in a screening process.

SQ Sequence 114 BP; 5 A; 10 C; 0 G; 0 T; 0 Indels 0; Gaps 0;

Query Match 9.7%; Score 29; DB 12; Length 114;
Best Local Similarity 5.8%; Pred. No. 5.60e-05;
Matches 6; Conservative 28; Mismatches 69; Indels 0; Gaps 0;

Db 3 bnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnnn 62
Cp 200 cacttgatcttcggtgtattcacacatctgtgtggcagggccgggttcattcca 141
63 bnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnnn 105
Cp 140 cacttcattcacatcttgactgaggttttcggtgtatctctg 98

RESULT 10
ID Q70466 standard; DNA; 114 BP.
AC Q70466;
DT 05-APR-1995 (first entry)
DE Generic DNA sequence to generate a random TSAR-9 peptide library.
KW TSAR; totally synthetic affinity reagent; synthetic; binding domain;
KW effector domain; concatenated heterofunctional protein; linker;
KW direct; rapid; detection; screening; treatment; generic; ss.
OS Synthetic.
FH Key Location/Qualifiers
FT misc_feature 55..60
FT /*tag= a
FT /note= "this sequence represents 'Z'; Z can be a
FT sequence of 6, 9 or 12 nucleotides (see
FT comments)"
PN W09418318-A.
PD 18-AUG-1994.
PF 01-FEB-1994; U00977.
PR 01-FEB-1993; US-013416.
PR 30-DEC-1993; US-176500.
PR 31-JAN-1994; US-189331.
PA (UYN-) UNIV NORTH CAROLINA.
PI Fowlkes DM, Kay BK;
P: WPI: 94-279739/34.
P-PSDB: R65152.
PT Identifying proteins or peptide(s) which bind a ligand - by
PT screening a recombinant vector library expressing fusion proteins
PT comprising a binding domain and an effector domain
PS Disclosure; Page 35; 255pp; English.
CC Q70466 is a generic DNA sequence used to generate random TSAR (Totally
CC Synthetic Affinity Reagents) peptides. This generic formula can also be
CC represented as follows: X(NNB)1(TGC)(NNB)10(TGC)2(NNB)42(NNB)8(TGC)(NNB)
CC -9X. X and Y are flanking restriction sites (X is not the same as Y)
CC that are not specified further. Other generic sequences are shown in
CC Q70466-68. Other specific peptides generated by these generic sequences
CC are shown in R65151-54. TSARs are concatenated heterofunctional proteins
CC or peptides, comprising at least two functional regions - a binding
CC domain with affinity for a ligand and a second effector peptide portion
CC that is chemically or biologically active. They may further comprise a
CC linker peptide between the 2 domains. The oligonucleotides are also
CC designed so that the expressed peptide contains 2 or 4 cysteine residues
CC positioned in, or flanking, the unpredicted or variant residues. These
CC residues confer some degree of conformational rigidity to the peptides.
CC The TSARs or comps. comprising a TSAR binding domain can be used in
CC vivo to deliver a chemically or biologically active moiety, eg. metal
CC ion, radioisotope, peptide, toxin or enzyme, to the specific target or
CC on the cell. They can also replace the function of macromolecules, eg.
CC monoclonal or polyclonal antibodies and therefore circumvent the need for
CC complex methods of hybridoma formation or in vivo antibody production.
CC The TSARs are easily characterised and have designed activity allowing
CC direct and rapid detection in a screening process.

SQ Sequence 114 BP; 0 A; 4 C; 4 G; 4 T;
Query Match 9.7%; Score 29; DB 12; Length 114;
Best Local Similarity 6.4%; Pred. No. 5.60e-05;
Matches 7; Conservative 28; Mismatches 74; Indels 0; Gaps 0;

Db 6 cnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnnn 65
Cp 122 ctgcaggtttccggtgtatctgcgaagcatctgtttgtgtccacgcactca 63

Db 66 bnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnnn 114
Cp 62 ccaacttaccatccaggttcgcatgtatgttcacagactccttctgt 14

RESULT 11
ID Q70465 standard; DNA; 114 BP.
AC Q70465;
DT 05-APR-1995 (first entry)
DE Generic DNA sequence to generate a random TSAR peptide library.
KW TSAR; totally synthetic affinity reagent; synthetic; binding domain;
KW effector domain; concatenated heterofunctional protein; linker;
KW direct; rapid; detection; screening; treatment; generic; ss.
OS Synthetic.
FH Key Location/Qualifiers
FT misc_feature 55..60
FT /*tag= a
FT /note= "this sequence represents 'Z'; Z can be a
FT sequence of 6, 9 or 12 nucleotides (see
FT comments)"
PN W09418318-A.
PD 18-AUG-1994.
PF 01-FEB-1994; U00977.
PR 01-FEB-1993; US-013416.
PR 30-DEC-1993; US-176500.
PR 31-JAN-1994; US-189331.
PA (UYN-) UNIV NORTH CAROLINA.
PI Fowlkes DM, Kay BK;
P: WPI: 94-279739/34.
P-PSDB: R65150 and R65151.
PT Identifying proteins or peptide(s) which bind a ligand - by
PT screening a recombinant vector library expressing fusion proteins
PT comprising a binding domain and an effector domain
PS Disclosure; Page 35; 255pp; English.
CC Q70465 is a generic DNA sequence used to generate random TSAR (Totally
CC Synthetic Affinity Reagents) peptides. This generic formula can also be
CC represented as follows: X(NNB)6(TGC)(NNB)112(TGC)14(TGC)(NNB)3X. X
CC and Y are flanking restriction sites (X is not the same as Y) that are
CC not specified further. Other generic sequences are shown in Q70466-68.
CC Other specific peptides generated by these generic sequences are shown in
CC R65151-54. TSARs are concatenated heterofunctional proteins or peptides,
CC comprising at least two functional regions - a binding domain with
CC affinity for a ligand and a second effector peptide portion that is
CC chemically or biologically active. They may further comprise a linker
CC peptide between the 2 domains. The oligonucleotides are also designed so
CC that the expressed peptide contains 2 or 4 cysteine residues positioned
CC in, or flanking, the unpredicted or variant residues. These residues
CC confer some degree of conformational rigidity to the peptides. The TSARs
CC or comps. comprising a TSAR binding domain can be used in vivo to
CC deliver a chemically or biologically active moiety, eg. metal ion,
CC radioisotope, peptide, toxin or enzyme, to the specific target or on the
CC cell. They can also replace the function of macromolecules, eg.
CC monoclonal or polyclonal antibodies and therefore circumvent the need
CC for complex methods of hybridoma formation or in vivo antibody
CC production. The TSARs are easily characterised and have designed
CC activity allowing direct and rapid detection in a screening process.
CC Sequence 114 BP; 0 A; 2 C; 2 G; 2 T;

Query Match 9.0%; Score 27; DB 12; Length 114;
Best Local Similarity 3.6%; Pred. No. 7.85e-04;
Matches 4; Conservative 29; Mismatches 79; Indels 0; Gaps 0;

Db 3 bnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnnn 62

	RESULT	14
ID	Q70467	standard; DNA; 114 BP.
AC	Q70467;	
DE	05-APR-1995	(first entry)
DT	Generic DNA sequence to generate a random TSAR peptide library.	
DD	TSAR; totally synthetic affinity reagent; synthetic binding domain;	
KW	TSAR; totally synthetic affinity reagent; synthetic binding domain;	

WILEY

(TM)

Release 3.1A John F. Collins, Biocomputing Research Unit.
Copyright (c) 1993-1998 University of Edinburgh, U.K.
Distribution rights by Oxford Molecular Ltd

psrch_nn n.a. - n.a. database search, using Smith-Waterman algorithm
Run on: Sat Apr 24 22:17:30 1999; MasPar time 484.95 Seconds
1107.807 Million cell updates/sec
Tabular output not generated.

Title: >US-08-968-800-1
Description: (1-300) from US08968800.seq
Perfect Score: 299
N.A. Sequence: 1 ggcgtggaagaaacagca.....atagctgtgaagacacagaa 300
Comp: cccgacctttcttctgtctt.....tatcgacattctgtctt

Scoring table: TABLE default
Gap 6

Nmatch STD : Dbase 0; Query 0

Searched: 2275026 seqs, 895388244 bases x 2

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database: embl-est56
1:em_est1 2:em_gssl 3:em_gss2 4:em_gss3
genbank-est109
5:gb_est1 6:gb_est10 7:gb_est11 8:gb_est12 9:gb_est13
10:gb_est14 11:gb_est15 12:gb_est16 13:gb_est17
14:gb_est18 15:gb_est19 16:gb_est2 17:gb_est20
18:gb_est21 19:gb_est3 20:gb_est4 21:gb_est5 22:gb_est6
23:gb_est7 24:gb_est8 25:gb_est9 26:gb_gssl 27:gb_gss2
28:gb_gss3 29:gb_gss4

Statistics: Mean 9.757; Variance 1.776; scale 5.495

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES				
Result No.	Score	Query Match	Description	Pred. No.
1	113	37.8	420 18 AA913032	0134a06.s1 Soares_NFL_ 1.05e-185
2	54	18.1	252 12 AA754459	97SN1787 Rice Immature 9.83e-62
3	52	17.4	252 12 AA754459	97SN1787 Rice Immature 8.07e-58
4	43	14.4	247 12 AA754458	97SN1784 Rice Immature 1.16e-40
5	41	13.7	247 12 AA754458	97SN1784 Rice Immature 5.67e-37
6	39	13.0	334 10 AA614649	np48a02.s1 NCI_CGAP.Br 2.46e-33
7	27	9.0	2275 11 AF034173	Homo sapiens ntcon2 co 4.91e-13
8	26	8.7	311 9 W09270	ma05h07.r1 Soares mous 1.68e-11
9	26	8.7	365 21 W82677	me98h07.r1 Soares mous 1.68e-11
10	26	8.7	473 21 AA028399	mi21e01.r1 Soares mous 1.68e-11
11	26	8.7	483 7 AA497481	vh29c03.r1 Soares mous 1.68e-11
12	26	8.7	554 12 AA793395	vn52d12.r1 Barstead mo 1.68e-11
13	24	8.0	2275 11 AF034173	Homo sapiens ntcon2 co 1.54e-08

RESULT LOCUS DEFINITION	1	AA913032	420 bp	mrna	EST	24-SEP-1998
ol34a06.s1 Soares_NFL.T.GBC.S1 Homo sapiens cDNA clone IMAGE:1525330 3' similar to SW:FBNI_MOUSE Q61554 FIBRILLIN 1						
PRECUSOR. ;, mRNA sequence.						
AA913032						
NID						
KEYWORDS						
SOURCE						
ORGANISM						
REFERENCE						
AUTHORS						
TITLE						
JOURNAL						
COMMENT						

ALIGNMENTS

Contact: Robert Strausberg, Ph.D.
Tel: (301) 496-1550
Email: Robert.Strausberg@nih.gov
This clone is available royalty-free through LNL; contact the IMAGE Consortium (info@image.llnl.gov) for further information.
Insert Length: 1169 Std Error: 0.00
Seq primer: -40ml3 fwd. ET from Amersham
High quality sequence stop: 68.
Location/Qualifiers
1. .420
/organism="Homo sapiens"
/note="Organ: pooled; Vector: pT73D-Pac (Pharmacia) with a modified polylinker; Site_1: Not I; Site_2: Eco RI; Equal amounts of plasmid DNA from three normalized libraries (fetal lung NBHL19w, testis NHT, and B-cell NCI-CGAP GCBI) were mixed, and ss circles were made in vitro. Following HAP purification, this DNA was used as tracer in a subtractive hybridization reaction. The driver was PCR-amplified cDNAs from pools of 5,000 clones made


```

/dev_stage="5 days after pollination"
/lab_host="E. coli SOLR"
BASE COUNT      5 a      21 c      35 t      179 others
ORIGIN

Query Match      17.4%; Score 52; DB 12; Length 252;
Best Local Similarity 12.8%; Pred. No. 8.07e-58;
Matches 27; Conservative 104; Mismatches 77; Indels 3; Gaps 3;

Db 27 GNBVVCVASHGNTVHNCBTRGTHCDCKNVNWSMTWTGTVNBNVSNVBNVBNVBNK 86
QY 2 gctggagagaagacagcaaggagctgtgaaagtacatgcgaacctgagtgatgttg 61

Db 87 VDVGNHTRCSWRBVTBMAHYDHTNCBBYNNNDYHMMHBBYBTGCMCTTMMCBWBYN 146
QY 62 gtgagt-gcgtgggaccaaacaatgcagatgcttccagataca-ccgggaaaacctg 119

147 TKCTASGHTSTNDVKSSTWTGVTBSYDKSMHGWCSBBVKYTKVSTTRATRSYTCV 206
120 cagtcaagatgtgaatgagtggaatgaaccccgccatgccacaaca-gatgtgtga 178

Db 207 RKVCVMWTKVKVKKYHVVBGCHTDSKCK 237
QY 179 atacacaggaagctacaagtgttttgcct 209

RESULT 4
LOCUS      AA754458      247 bp      mRNA      EST      20-JAN-1998
DEFINITION 97SN1784 Rice Immature Seed Lambda ZAPII cDNA Library Oryza sativa
ACCESSION  AA754458
NID        G2801164
KEYWORDS   EST.
SOURCE     rice.
ORGANISM   Oryza sativa
            Eukaryotae; Viridiplantae; Charophyta/Embryophyta group;
            Embryophyta; Tracheophyta; seed plants; Magnoliophyta; Liliopsida;
            Poales; Poaceae; Oryza.
REFERENCE  1 (bases 1 to 247)
AUTHORS   Nahm,B.H., Kim,J.K., Cheong,J.J., Kim,S.I., Hahn,T.R., Moon,E.P.,
            Kim,W.T., Kim,W.Y., Yang,M.S., Park,R.D., Sohn,U.I., Kang,K.Y.,
            Lee,M.C. and Eun,M.Y.
            Large-scale Sequencing Analysis of ESTs from Rice Immature Seed
            Unpublished (1998)
TITLE     Large-scale Sequencing Analysis of ESTs from Rice Immature Seed
JOURNAL   Unpublished (1998)
COMMENT   Contact: Eun M.Y.
            Department of Cytoogenetics
            National Inst. of Agri. Sci. and Tech, RDA
            Suwon, Kyunggido, Korea
            Tel: 82 331 290 0301
            Fax: 82 331 290 0307
            Email: myeun@sun20.astl.re.kr
            Submitted by Baek Hie Nahm, Dept of Biological Science, Myongji
            University, Yongin, Korea. 449-728 bhnahebio@server.myongji.ac.kr
            Seq primer: M13 Reverse Primer.
            Location/Qualifiers
            1. .247
            /organism="Oryza sativa"
            /cultivar="Milyang23"
            /note="Vector: pBluescript SK(+); Site_1: EcoRI; Site_2:
            XhoI; Directional cDNA library inserted into lambda ZAPII
            vector at 5' end with EcoRI and 3' end with Xho I site."
            /db_xref="taxon:4530"
            /clone="97SN1784"
            /tissue_type="Immature Seed"
            /dev_stage="5 days after pollination"
            /lab_host="E. coli SOLR"

BASE COUNT      7 a      16 c      21 g      34 t      169 others
ORIGIN

Query Match      14.4%; Score 43; DB 12; Length 247;
Best Local Similarity 15.5%; Pred. No. 1.16e-40;
Matches 18; Conservative 87; Mismatches 72; Indels 1; Gaps 1;

Db 53 WNCSDNAHCRTYBMYARSKYGTBYYSWNVDINTGGTGVGKTTVNVHSGWNNRCSN 112
QY 24 agtctgtgaagtacatgcgaacctg-gatgtaagtgttgtagtcggtgggaccaaaca 82

Db 113 SVVYVBTAYCDYBHYBDRANHVDTCTNDRGVCNNTASDNGTSATKRVGTGDKTSDC 172

```

```

Matches 37; Conservative 98; Mismatches 101; Indels 2; Gaps 2;

Db 1 HWDCTMTNTVWRGCCCBAMNKHHTHMTBWMCCVVRVGTFTTNGKHNGRITTTWDCSDNA 60
QY 260 catgctcctngagttcacacacagtagctggtgcatgagcagtggtgccc-actgagggcaaaa 202

Db 61 HCRITYBMYARSKYGTBYYSWNVDINTGGTGVGKTTVNVHSGWNNRCSNVSVVYVBT 120
QY 201 gcactgtagcttcgtgtattcacacatctgtgtgcatggtggtggtggttccattcc 142

Db 121 AYCDYBHYBDRANHVDTCTNDRGVCNNTASDNGTSATKRVGTGDKTSDCGGCGWRKV 180
QY 141 acactcatcatcattgtagtcgaggtttcccggtgtatctcgtgaaagcattgctgatt 82

Db 181 TYGSSBYRRCGVNVMVTRTSMWTDKSTRK-BSDMRSRVRHYGRWMBKRGMSRNW 237
QY 81 gttgtgtccagcactcaccaactacattcaggttcgcatgtagcttccagact 24

RESULT 5
LOCUS      AA754458      247 bp      mRNA      EST      20-JAN-1998
DEFINITION 97SN1784 Rice Immature Seed Lambda ZAPII cDNA Library Oryza sativa
ACCESSION  AA754458
NID        G2801164
KEYWORDS   EST.
SOURCE     rice.
ORGANISM   Oryza sativa
            Eukaryotae; Viridiplantae; Charophyta/Embryophyta group;
            Embryophyta; Tracheophyta; seed plants; Magnoliophyta; Liliopsida;
            Poales; Poaceae; Oryza.
REFERENCE  1 (bases 1 to 247)
AUTHORS   Nahm,B.H., Kim,J.K., Cheong,J.J., Kim,S.I., Hahn,T.R., Moon,E.P.,
            Kim,W.T., Kim,W.Y., Yang,M.S., Park,R.D., Sohn,U.I., Kang,K.Y.,
            Lee,M.C. and Eun,M.Y.
            Large-scale Sequencing Analysis of ESTs from Rice Immature Seed
            Unpublished (1998)
TITLE     Large-scale Sequencing Analysis of ESTs from Rice Immature Seed
JOURNAL   Unpublished (1998)
COMMENT   Contact: Eun M.Y.
            Department of Cytoogenetics
            National Inst. of Agri. Sci. and Tech, RDA
            Suwon, Kyunggido, Korea
            Tel: 82 331 290 0301
            Fax: 82 331 290 0307
            Email: myeun@sun20.astl.re.kr
            Submitted by Baek Hie Nahm, Dept of Biological Science, Myongji
            University, Yongin, Korea. 449-728 bhnahebio@server.myongji.ac.kr
            Seq primer: M13 Reverse Primer.
            Location/Qualifiers
            1. .247
            /organism="Oryza sativa"
            /cultivar="Milyang23"
            /note="Vector: pBluescript SK(+); Site_1: EcoRI; Site_2:
            XhoI; Directional cDNA library inserted into lambda ZAPII
            vector at 5' end with EcoRI and 3' end with Xho I site."
            /db_xref="taxon:4530"
            /clone="97SN1784"
            /tissue_type="Immature Seed"
            /dev_stage="5 days after pollination"
            /lab_host="E. coli SOLR"

BASE COUNT      7 a      16 c      21 g      34 t      169 others
ORIGIN

Query Match      13.7%; Score 41; DB 12; Length 247;
Best Local Similarity 10.1%; Pred. No. 5.67e-37;
Matches 18; Conservative 87; Mismatches 72; Indels 1; Gaps 1;

Db 53 WNCSDNAHCRTYBMYARSKYGTBYYSWNVDINTGGTGVGKTTVNVHSGWNNRCSN 112
QY 24 agtctgtgaagtacatgcgaacctg-gatgtaagtgttgtagtcggtgggaccaaaca 82

Db 113 SVVYVBTAYCDYBHYBDRANHVDTCTNDRGVCNNTASDNGTSATKRVGTGDKTSDC 172

```

QY 83 aatgcagatgtcttcacagataccagcggaacacctcgcagtcgaagtgtg 142
 Db 173 GGCGRKVTYGSBYRCGVNVMVVRTSMWTDKSTKMSMDMSRRSRVHYGRWMBNKK 230
 QY 143 gaatgaaccccgccgcatgccaacagatgtgtgaatacacacggaagctacaagt 200

RESULT 6
 LOCUS AA614649 334 bp mRNA EST 16-OCT-1997
 DEFINITION np48a02.s1 NCI-CGAP_Br1.1 Homo sapiens cDNA clone IMAGE:1129514,
 mRNA sequence.
 ACCESSION AA614649
 NID 92466845
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryotae; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;
 Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 334)
 AUTHORS NCI-CGAP
 TITLE NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
 JOURNAL National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
 Tumor Gene Index
 COMMENT Unpublished (1997)
 Contact: Robert Strausberg, Ph.D.
 Tel: (301) 496-1550
 Email: Robert.Strausberg@nih.gov
 Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.
 Emmert-Buck, M.D., Ph.D.
 cDNA Library Preparation: M. Bento Soares, Ph.D.
 cDNA Library Arrayed by: Greg Lennon, Ph.D.
 DNA Sequencing by: Washington University Genome Sequencing Center
 Clone distribution: NCI-CGAP clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
 www-bio.llnl.gov/bbrp/image/image.html

Insert Length: 1982 Std Error: 0.00
 Seq primer: -40m3 fwd. Et from Amersham

High quality sequence stop: 318.
 Location/Qualifiers

FEATURES

source
 1..334
 /organism="Homo sapiens"
 /note="Vector: pT73D-Pac (Pharmacia) with a modified
 polylinker; 1st strand cDNA was prepared from pooled bulk
 breast tumor tissue, and was then primed with a Not I -
 oligo(dT) primer. Double-stranded cDNA was ligated to Eco
 RI adaptors (Pharmacia), digested with Not I and cloned
 into the Not I and Eco RI sites of the modified pT73
 vector. Library is not normalized. (The normalized
 version of this library is NCI-CGAP Br2.) Library was
 constructed by Bento Soares and M. Fatima Bonaldo."
 /db_xref="taxon:9606"
 /clone="IMAGE:1129514"
 /clone.lib="NCI-CGAP_Br1.1"
 /sex="female, pooled"
 /tissue_type="breast"
 /lab_host="DH10B"
 BASE COUNT 98 a 60 c 65 g 111 t
 ORIGIN

Query Match 13.0%; Score 39; DB 10; Length 334;
 Best Local Similarity 68.9%; Pred. No. 2.46e-33;
 Matches 71; Conservative 0; Mismatches 32; Indels 0; Gaps 0;

Db 144 TTCTACCTTGATACAGGTTTTCACGATACACGATGACATGCTGCTTTGGC 203
 Cp 134 ttcaacttgatgcaggtttcccggtgtatctggaagcatctgctattgttgg 75
 Db 204 CGATACATTCACATGTTTGCATCGGTGGCACACAGCTT 246
 Cp 74 cccagcactcaccacactatcatcaggttcgcattgactt 32

RESULT 7
 LOCUS AF034173 2275 bp mRNA EST 22-DEC-1997
 DEFINITION Homo sapiens ntcon2 contig mRNA, partial sequence.
 ACCESSION AF034173
 NID 92707735
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryotae; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;
 Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 2275)
 AUTHORS Tripodis, N. and Ragoussis, J.
 TITLE Generation of a transcription map in the region immediately
 centromeric to human MHC across the 6p21.2-6p21.3 chromosomal
 boundary
 JOURNAL Unpublished
 REFERENCE 2 (bases 1 to 2275)
 AUTHORS Tripodis, N. and Ragoussis, J.
 TITLE Direct Submission
 JOURNAL Submitted (13-NOV-1997) Division of Medical and Molecular Genetics,
 Guy's Hospital, 7th floor, Guy's Tower, London SE1 9RT, UK
 FEATURES
 Location/Qualifiers
 1..2275
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /chromosome="6"
 /map="6p21.3"
 /clone="ntcon2 contig"
 /tissue_type="liver; brain"
 /dev_stage="fetus"
 /note="similar to Br140"
 BASE COUNT 438 a 619 c 470 g 599 t 149 others
 ORIGIN

Query Match 9.0%; Score 27; DB 11; Length 2275;
 Best Local Similarity 10.6%; Pred. No. 4.91e-13;
 Matches 10; Conservative 50; Mismatches 34; Indels 0; Gaps 0;

Db 1476 AAMWRYKRWKRRGRKRRMTGMVKRYMMAMMAMCMMACWYKMGKKCKWKRYK 1535
 QY 164 aacacagatgctggaatacacaggaagctacagtgctttgctcagggccacatgc 223
 Db 1536 KYTSTYKWSRWYTTTYYWYWCWCTSMKSKAS 1569
 QY 224 tcacgcagatgctacgtgtgtgaactcnaggac 257

RESULT 8
 LOCUS W09270 311 bp mRNA EST 01-OCT-1997
 DEFINITION ma05h07.r1 Soares mouse p3NMF19.5 Mus musculus cDNA clone 303709 5',
 similar to SW.TLD.DROME P25723 DORSAL-VENTRAL PATTERNING TOLLOID
 PROTEIN PRECURSOR ;, mRNA sequence.
 ACCESSION W09270
 NID 91283586
 KEYWORDS EST.
 SOURCE house mouse.
 ORGANISM Mus musculus
 Eukaryotae; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;
 Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 REFERENCE 1 (bases 1 to 311)
 AUTHORS Marra, M., Hillier, L., Allen, M., Bowles, M., Dietrich, N., Dubuque, T.,
 Geisler, S., Kucaba, T., Lacy, M., Le, M., Martin, J., Morris, M.,
 Schellenberg, K., Steptoe, M., Tan, F., Underwood, K., Moore, B.,
 Theising, B., Wylie, T., Lennon, G., Soares, B., Wilson, R. and
 Waterston, R.
 TITLE The WashU-HMI Mouse EST Project
 JOURNAL Unpublished (1996)
 COMMENT Contact: Marra M/Mouse EST Project
 WashU-HMI Mouse EST Project
 Washington University School of Medicine
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108

Tel: 314 286 1800
Fax: 314 286 1810

Email: mouseest@watson.wustl.edu
This clone is available royalty-free through LNL ; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.

MGI: 200469

Possible reversed clone: similarity on wrong strand

Seq primer: mob.REGA+ET
High quality sequence stop: 290.

FEATURES

source

1. .311
/organism="Mus musculus"
/note="vector: pT73D (Pharmacia) with a modified
polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA
was primed with a Not I - oligo(dT) primer [5'
TGTTACCAATCTGAAGTGGAGCGCGCATTTTTTTTTTTT 3'],
double-stranded cDNA was size selected, ligated to Eco RI
adapters (Pharmacia), digested with Not I and cloned into
the Not I and Eco RI sites of a modified pT73 vector
(Pharmacia). Library went through one round of
normalization to a Cot = 5. Library constructed by Bento
Soares and M.Fatima Bonaldo. RNA was kindly provided by
Dr. Minoru Ko (Wayne State University)."
/db_xref="taxon:10090"
/clone="303709"
/clone_lib="Soares mouse p3NMF19.5"
/dev_stage="19.5 dpc total fetus"
/lab_host="DH10B (ampicillin resistant)"

BASE COUNT 79 a 82 c 88 g 62 t

ORIGIN

Query Match 8.7%; Score 26; DB 9; Length 311;
Best Local Similarity 72.4%; Pred. No. 1.68e-11;
Matches 42; Conservative 0; Mismatches 16; Indels 0; Gaps 0;

Db 20 ATGCAGCAGACTGTGTGAACAGCTTCGGCAGCTACGAGTGTCTAGTCCGCGAGTGGC 77

QY 159 atgccaacacagatgtgtgaatacacagcgagctacaagtgtttgtcctcagtgcc 216

RESULT 9

LOCUS W82677 365 bp mRNA EST 12-SEP-1996
DEFINITION me98d07.r1 Soares mouse p3NMF19.5 Mus musculus cDNA clone 403597 5',
similar to gb:M22488 BONE MORPHOGENETIC PROTEIN 1 PRECURSOR
(HUMAN);, mRNA sequence.

ACCESSION

W82677

g1539666

EST.

house mouse.

Mus musculus

Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata;

Vertebrata; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;

Mus.

1 (bases 1 to 365)

Marra, M., Hillier, L., Allen, M., Bowles, M., Dietrich, N., Dubuque, T.,

Geisel, S., Kucaba, T., Lacy, M., Le, M., Martin, J., Morris, M.,

Schellenberg, K., Steptoe, M., Tan, F., Underwood, K., Moore, B.,

Theising, B., Wylie, T., Lennon, G., Soares, B., Wilson, R. and

Waterston, R.

The WashU-HMI Mouse EST Project

Unpublished (1996)

On Sep 13, 1996 this sequence version replaced gi:1393691.

Contact: Marra M/Mouse EST Project

WashU-HMI Mouse EST Project

Washington University School of MedicineP

4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108

Tel: 314 286 1800

Fax: 314 286 1810

Email: mouseest@watson.wustl.edu

This clone is available royalty-free through LNL ; contact the

IMAGE Consortium (info@image.llnl.gov) for further information.

MGI:247365

Seq primer: -28M13 rev2 from Amersham
High quality sequence stop: 353.

FEATURES

source

1. .365

/organism="Mus musculus"

/note="vector: pT73D (Pharmacia) with a modified

polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA

was primed with a Not I - oligo(dT) primer [5'

TGTTACCAATCTGAAGTGGAGCGCGCATTTTTTTTTTTT 3'],

double-stranded cDNA was size selected, ligated to Eco RI

adapters (Pharmacia), digested with Not I and cloned into

the Not I and Eco RI sites of a modified pT73 vector

(Pharmacia). Library went through one round of

normalization to a Cot = 5. Library constructed by Bento

Soares and M.Fatima Bonaldo. RNA was kindly provided by

Dr. Minoru Ko (Wayne State University)."

/db_xref="taxon:10090"

/clone="403597"

/clone_lib="Soares mouse p3NMF19.5"

/dev_stage="19.5 dpc total fetus"

/lab_host="DH10B (ampicillin resistant)"

<1. .>365

BASE COUNT 101 a 97 c 95 g 72 t

ORIGIN

Query Match 8.7%; Score 26; DB 21; Length 365;
Best Local Similarity 72.4%; Pred. No. 1.68e-11;
Matches 42; Conservative 0; Mismatches 16; Indels 0; Gaps 0;

Db 177 ATGCAGCAGACTGTGTGAACAGCTTCGGCAGCTACGAGTGTCTAGTCCGCGAGTGGC 234

QY 159 atgccaacacagatgtgtgaatacacagcgagctacaagtgtttgtcctcagtgcc 216

RESULT 10

LOCUS AA028399 473 bp mRNA EST 16-AUG-1996
DEFINITION mi2le01.r1 Soares mouse embryo NBME13.5 14.5 Mus musculus cDNA
clone 464184 5' similar to SW:TLD.DROME P25723 DORSAL-VENTRAL
PATTERNING TOLLOID PROTEIN PRECURSOR ;, mRNA sequence.

ACCESSION

AA028399

91494520

EST.

house mouse.

Mus musculus

Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata;

Vertebrata; Eutheria; Rodentia; Sciurognathi; Myomorpha; Muridae;

Murinae; Mus.

1 (bases 1 to 473)

Marra, M., Hillier, L., Allen, M., Bowles, M., Dietrich, N., Dubuque, T.,

Geisel, S., Kucaba, T., Lacy, M., Le, M., Martin, J., Morris, M.,

Schellenberg, K., Steptoe, M., Tan, F., Underwood, K., Moore, B.,

Theising, B., Wylie, T., Lennon, G., Soares, B., Wilson, R. and

Waterston, R.

The WashU-HMI Mouse EST Project

Unpublished (1996)

Contact: Marra M/Mouse EST Project

WashU-HMI Mouse EST Project

Washington University School of MedicineP

4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108

Tel: 314 286 1800

Fax: 314 286 1810

Email: mouseest@watson.wustl.edu

This clone is available royalty-free through LNL ; contact the

IMAGE Consortium (info@image.llnl.gov) for further information.

MGI:278000

Seq primer: -28M13 rev2 from Amersham

High quality sequence stop: 462.

1. .473

/organism="Mus musculus"

/strain="C57BL/6J"

/note="vector: pT73D-Pac (Pharmacia) with a modified

```
Seq primer: -28m13 rev2 Et from Amersham
High quality sequence stop: 420.
Location/Qualifiers
1. .483
/organism="Mus musculus"
/strain="C57BL/6J"
/note="Organ: mammary gland; Vector: pT73D-Pac
(Pharmacia) with a modified polylinker; Site.1:
Site.2: Eco RI; 1st strand cDNA was primed with a Not I -
oligo(dT) primer [5',
```

FEATURES source

```

/organism="Mus musculus"
/strain="C3H"
/note="Vector: pT7T3D-Pac (Pharmacia) with a modified
polylinker; Site_1: EcoRI; Site_2: NotI; 1st strand cDNA
was primed with a Not I - oligo(dT) primer 15',
TGTTACGAATCTGAATGGAGCGCCGCCCCCTTTTTTTTTTTTTTTT
3']; Double-stranded cDNA was ligated to Eco RI adaptors
[AAATTCGATCCTTC], digested with Not I and cloned into the
Not I and Eco RI sites of the modified pT7T3 vector.
Library constructed by Bob Barstead. The C2C12 cell line
(available from ATCC, catalog # CRL-1772) differentiates

```


Deuterostomia; Chordata; Vertebrata; Gnathostomata; Osteichthyes;
Sarcopterygii; Chonata; Tetrapoda; Amniota; Mammalia; Theria;
Eutheria; Archonta; Primates; Catarrhini; Hominoidea; Homo.
1 (bases 1 to 446)

REFERENCE
AUTHORS

Hillier, L., Clark, N., Dubuque, T., Elliston, K., Hawkins, M.,
Holman, M., Hultman, M., Kucaba, T., Le, M., Lennon, G., Marra, M.,
Parsons, J., Rifkin, L., Rohlfing, T., Soares, M., Tan, F.,
Trevaskis, E., Waterston, R., Williamson, A., Wohlmann, P. and
Wilson, R.
The WashU-Merck EST Project
Unpublished (1995)

TITLE
JOURNAL
COMMENT

Contact: Wilson RK
WashU-Merck EST Project
Washington University School of Medicine
444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu

High quality sequence stops: 326
Source: IMAGE Consortium, LLNL

This clone is available royalty-free through LLNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.

FEATURES
source

1..446
/organism="Homo sapiens"
/clone="201021"

BASE COUNT
ORIGIN

138 a 93 c 104 g 107 t 4 others

Query Match
Best Local Similarity 7.7%; Score 23; DB 16; Length 446;
Matches 35; Conservative 0; Mismatches 12; Indels 0; Gaps 0;

Db 121 AGATGTGTGAACACATTGGAGCTACTACTGCAATGTCACATGG 167
|||||
QY 169 agatgtgtgaatacacacaggaagctacaagtgtttgcctcagtg 215

Search completed: Sat Apr 24 22:25:49 1999
Job time : 499 secs.

WIPESHIELD (TM)

Release 3.1A John F. Collins, Biocomputing Research Unit.
Copyright (c) 1993-1998 University of Edinburgh, U.K.
Distribution rights by Oxford Molecular Ltd

Search_nnn n.a. - n.a. database search, using Smith-Waterman algorithm
on: Thu Apr 29 03:05:32 1999; MasPar time 2494.84 Seconds
1525.983 Million cell updates/sec
Tabular output not generated.
Title: >US-08-968-800-2
Description: (1-1611) from US08968800.seq
Perfect Score: 1607
N.A. Sequence: 1 ggcgtgagaagaacagca.....aattactagctgaaaaattg 1611
Comp: ccgacctttttgttcgtt.....ttaagatcgacttttaac

Scoring table: TABLE default
Gap 6
Nmatch STD : Dbase 0; Query 0
Searched: 602357 seqs, 1181590623 bases x 2
Post-processing: Minimum Match 0%
Listing first 45 summaries
Database: emb157
1:em_ba 2:em_fun 3:em_htg 4:em_hum1 5:em_hum2 6:em_in
7:em_or 8:em_ov 9:em_ov 10:em_pat 11:em_ph 12:em_pl
13:em_ro 14:em_un 15:em_v1
Database: genbank110
16:gb_ba1 17:gb_ba2 18:gb_htg 19:gb_in 20:gb_om 21:gb_ov
22:gb_pat 23:gb_ph 24:gb_pl 25:gb_pi2 26:gb_pl1
27:gb_pr2 28:gb_pr3 29:gb_ro 30:gb_st 31:gb_sts 32:gb_sy
33:gb_un 34:gb_v1
Statistics: Mean 11.423; Variance 5.514; scale 2.072

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES							
Result No.	Score	Query Match	Length	DB	ID	Description	Pred. No.
1	108	6.7	7218	22	I66494	Sequence 14 from patent	1.81e-60
2	41	2.6	215	22	I28278	Sequence 5 from patent	1.24e-09
3	37	2.3	2070	19	AF051401	Caenorhabditis elegans	4.99e-07
4	37	2.3	2139	19	AF051402	Caenorhabditis elegans	4.99e-07
5	37	2.3	10772	19	AF012089	Drosophila melanogaste	4.99e-07
6	35	2.2	7218	22	I66494	Sequence 14 from patent	9.13e-06
7	36	2.2	9662	26	HUMFIBRLN	Homo sapiens fibrillin	2.15e-06
8	36	2.2	9940	26	HSFIBMR	H.sapiens mRNA for fib	2.15e-06
9	35	2.2	10772	19	AF012089	Drosophila melanogaste	9.13e-06
10	36	2.2	74371	27	AC005369	Homo sapiens chromosom	2.15e-06
11	33	2.1	2156	26	HSFIBUC	H.sapiens mRNA for fib	1.56e-04
12	33	2.1	2349	26	HSFIBUA	H.sapiens mRNA for fib	1.56e-04
13	33	2.1	2359	26	U01244	Human fibrulin-1D mRNA,	1.56e-04

14	33	2.1	2525	26	HSFIBUB H.sapiens mRNA for fib	1.56e-04
15	32	2.0	9830	29	MMU22493 Mus musculus fibrillin	6.25e-04
16	32	2.0	11724	29	AF007248 Mouse fibrillin mutant f1	6.25e-04
17	32	2.0	11953	29	MUSFBN1A Mouse fibrillin (Fbn-1	6.25e-04
18	30	1.9	215	22	I28278 Sequence 5 from patent	9.42e-03
19	30	1.9	965	22	AR024229 Sequence 22 from patent	9.42e-03
20	31	1.9	1561	26	HUMCLNXX Human calnexin mRNA, c	2.45e-03
21	31	1.9	1288	26	HUMCALNEXI Human calnexin mRNA, c	2.45e-03
22	31	1.9	3881	26	HUMCALIEF Human calnexin integral	2.45e-03
23	31	1.9	4117	26	HUMIP90 Homo sapiens Chromosom	9.42e-03
24	30	1.9	74371	27	AC005369 Homo sapiens Chromosom	9.42e-03
25	29	1.8	151	29	MMNIDEX12 M.musculus nid gene (e	3.53e-02
26	29	1.8	985	22	AR024229 Sequence 22 from patent	3.53e-02
27	29	1.8	2358	21	AF051400 Gallus gallus fibrulin-	3.53e-02
28	29	1.8	2709	29	MMB908 M.musculus (isolate MK	3.53e-02
29	29	1.8	3346	21	AF051399 Gallus gallus fibrulin-	3.53e-02
30	29	1.8	4171	29	MMNIDOG Mouse mRNA for nidogen	3.53e-02
31	29	1.8	4633	21	AF027596 Danio rerio tolloid mr	3.53e-02
32	29	1.8	5959	29	MMENACT Mouse mRNA for entacti	3.53e-02
33	29	1.8	9859	29	MUSFBN2 Mus musculus fibrillin	3.53e-02
34	29	1.8	10172	26	HSU03272 Human fibrillin-2 mRNA	3.53e-02
35	29	1.8	216021	27	HUAC004787 Homo sapiens Chromosom	3.53e-02
36	29	1.8	238893	18	HS90L6 Human DNA sequence ***	3.53e-02
37	28	1.7	1548	24	GMNGM16 Soybean Ngm-16 gene co	1.29e-01
38	28	1.7	2823	26	HSFIB5 H.sapiens mRNA for fib	1.29e-01
39	28	1.7	8950	20	BOVXAAA Bos taurus mRNA, compl	1.29e-01
40	27	1.7	35593	19	CELC56E5 Caenorhabditis elegans	4.60e-01
41	27	1.7	37001	27	AC005513 Homo sapiens Chromosom	4.60e-01
42	27	1.7	125536	18	AC003021 *** SEQUENCING IN PROG	4.60e-01
43	28	1.7	158427	27	AC002553 Homo sapiens Chromosom	1.29e-01
44	27	1.7	183338	18	HS329A5 Human DNA sequence ***	4.60e-01
45	27	1.7	216021	27	HUAC004787 Homo sapiens Chromosom	4.60e-01

ALIGNMENTS

RESULT 1	I66494	7218 bp	DNA	PAT	23-DEC-1997
LOCUS	Sequence 14 from patent US 5670367.				
DEFINITION	166494				
ACCESSION	166494				
NID	92724471				
KEYWORDS	Unknown.				
SOURCE	Unknown.				
ORGANISM	Unclassified.				
REFERENCE	1 (bases 1 to 7218)				
AUTHORS	Dorner, F., Scheiflinger, F. and Falkner, F. Gunter.				
TITLE	Recombinant fowlpox virus				
JOURNAL	Patent: US 5670367-A 14 23-SEP-1997;				
FEATURES	Location/Qualifiers				
source	1..7218				
BASE COUNT	1944 a 1491 c 1486 g 1929 t				
ORIGIN	/organism="unknown"				
Query Match 6.7%; Score 108; DB 22; Length 7218;					
Best Local Similarity 1.6%; Pred. No. 1.81e-60;					
Matches 6; Conservative 231; Mismatches 129; Indels 0; Gaps 0;					
Db	1075	Y	Y	Y	Y
Cp	1031	t	t	t	t
Db	1135	Y	Y	Y	Y
Cp	971	c	c	c	c
Db	1195	Y	Y	Y	Y
Cp	911	t	t	t	t
Db	1255	Y	Y	Y	Y

Cp 851 ttccctttttacccctcagagagttccgcctctggaacatctctcttcattagtgag 792
Db 1315 YY 1374
Cp 791 ggcgcagctccctcaggtaggagtcctggtggtctcgggtaacatttttaatt 732
Db 1375 YY 1434
Cp 731 ttgcctcttttctgctgttttctgagcaagcaactctgtattctgtcttgatg 672
Db 1435 GTACCA 1440
Cp 671 gtacca 666

RESULT 2 I28278 215 bp DNA PAT 30-OCT-1996
LOCUS Sequence 5 from patent US 5569830.
DEFINITION I28278
ACCESSION g1819054

WORDS
SOURCE Unknown.
ORGANISM Unknown.

REFERENCE 1 (bases 1 to 215)

AUTHORS Bennett, A., Labavitch, J.M., Powell, A. and Stotz, H.
TITLE Plant inhibitors of fungal polygalacturonases and their use to control fungal disease

JOURNAL Patent: US 5569830-A 5 29-OCT-1996;
FEATURES Location/Qualifiers
1. 215

BASE COUNT 15 a 8 c 25 g 26 t 141 others
ORIGIN /organism="unknown"

Query Match 2.6%; Score 41; DB 22; Length 215;
Best Local Similarity 16.3%; Pred. No. 1.24e-09;
Matches 34; Conservative 82; Mismatches 91; Indels 2; Gaps 2;

Db 7 SSSSVVSRASCDNDKAKKGGNTTSSWTDCCNRRTGWVCDTDTTVRVNNDGHNKYSANY 66
Cp 343 ccaggcgagctcctgagtgagcacagcagctggtggtctctctctctcagc 284

Db 67 NYGGNNVGAAKTHYHTVNVSGADSKVTVDSYNASGTSNGTGNRSGADSYGSKTA 126
Cp 283 tatactgacagttatcatggca-catgctctngagttcacacagctagcatctggcag 225

Db 127 MTSNRRTGTANNVDSRNNGDASVGSNDKNTKKHAKNSADGKVGSKNNGDRNNRYGTGK 186
Cp 224 agcatgtggccactggaggcaagacactgttagctccgtgtattc-acacatctgtg 166

Db 187 SNVSNCGGNRDRVSSYANNKCGSSCT 215
Cp 165 ttggcatgcccggggtttcattccacact 137

RESULT 3 AF051401 2070 bp mRNA INV 02-AUG-1998
LOCUS Caenorhabditis elegans fibulin-1 isoform D precursor (FBLN1) mRNA,
DEFINITION alternatively spliced product, complete cds.
ACCESSION AF051401
NID G3372520

WORDS
SOURCE Caenorhabditis elegans.
ORGANISM Caenorhabditis elegans

REFERENCE 1 (bases 1 to 2070)

AUTHORS Barth, J.L., Argraves, K.M., Roark, E.F., Little, C.D. and

TITLE Isolation of chicken and nematode fibulin-1 homologs and
JOURNAL characterization of the nematode fibulin-1 gene
Unpublished

REFERENCE
AUTHORS

Barth, J.L., Argraves, K.M., Roark, E.F., Little, C.D. and
Argraves, W.S.

TITLE Direct Submission

JOURNAL Submitted (26-FEB-1998) Cell Biology and Anatomy, Medical
University of South Carolina, 171 Ashley Avenue, Charleston, SC
29401, USA

FEATURES Location/Qualifiers
1. 2070

source /organism="Caenorhabditis elegans"
/db_xref="taxon:6239"
/chromosome="IV"

gene /map="between elt-1 and daf-10"

CDS 1. 2070

/gene="FBLN1"

/gene="FBLN1"

/note="extracellular matrix protein"

/codon_start=1

/product="fibulin-1 isoform D precursor"

/translation="MRICLLLAFLVAETFANELTRCAGGTRHFKNSNTSSIKSEG
TSMTCORAAISCLRLDNDACDSTDAKEEPCSPNIIIGGLKKECCDCLLAK
DILNNPECPVAPVGFSAQLRSFNKCCNGDIEITHASEIITGRPLNDPHVLGDRCA
SSHCEHLCHDRGKEKVECSRSRGLADPGMACVDIDECATLMDCCLESQRLNTPGS
FKCIRTLCSTGYAMDSETRCROVDENGLSHGCGPLQCRNTOGSTRCDAAKKGDD
ELQNPMTGECTSITCPNGYYPKNGMNDIDECVTGHNCGAGGECVNTGSPRCQKGN
LCANGYEVNGATGCEDEVNECOQGVGSGMECINLPGYKCKGPGYFNDAKKECDV
DECIFAGHVCDSALSCINLTIGSEFCKCPQFQASDGRCEDEYECTGTAAACEKRC
VNIFSYQICDRGLFADPGDKEDIDECISWAGSNDLCMGCIINTKSYLQCCPP
GYKIOPDGRICVDDEAGCEAGSKVCVNTLGSFKCHSDCFTNTHDSLKNQIA
DGYSIKVCSTEDTECLGNTHREVLVQFRAVPSLKTILISPIEVSRIVTHMGVPSVDY
NLDTVGORHFRIVQERNIGIVQLVKPISGPTVETIKVNIHTKSTGVLAFNEALIEI
SVSKYPT"

BASE COUNT 640 a 337 c 493 g 600 t
ORIGIN

Query Match 2.3%; Score 37; DB 19; Length 2070;

Best Local Similarity 65.3%; Pred. No. 4.99e-07;

Matches 79; Conservative 0; Mismatches 42; Indels 0; Gaps 0;

Db 544 TCTGGATTCGACTTGGCACCAGATGCGTGTGTAGATATTGATGTGTCACACA 603
Qy 328 tcaggactccgcctggcccaaatggagagactgtctagatattgatgtgcctct 387

Db 604 CTGATGACCATGTCTTGGATCTCAACGATGTTGAATCTCTCTGGAAGCTTTAAATGC 663
Qy 388 ggtaaagtcactgtcctcaatcgaagtgtgcaacattggaagactactgc 447

Db 664 A 664

Qy 448 a 448

RESULT 4 AF051402 2139 bp mRNA INV 02-AUG-1998

LOCUS Caenorhabditis elegans fibulin-1 isoform C precursor (FBLN1) mRNA,
DEFINITION alternatively spliced product, complete cds.

ACCESSION AF051402

NID G3372523

WORDS

SOURCE Caenorhabditis elegans.

ORGANISM Caenorhabditis elegans

REFERENCE 1 (bases 1 to 2139)

AUTHORS Barth, J.L., Argraves, K.M., Roark, E.F., Little, C.D. and

TITLE Isolation of chicken and nematode fibulin-1 homologs and
JOURNAL characterization of the nematode fibulin-1 gene
Unpublished

REFERENCE 2 (bases 1 to 2139)


```
Matches 20; Conservative 63; Mismatches 46; Indels 0; Gaps 0;
```

[illegible]

RESULT	6				
LOCUS					
DEFINITION			7218 bp	DNA	
ACCESSION			Sequence 14	from patent US 5670357.	PAT
			I66494		
			I66494		
			92724471		
					23-DEC-1997

WORDS
 SOURCE
 ORGANISM
 REFERENCE
 AUTHORS
 TITLE
 JOURNAL
 FEATURES
 source

BASE COUNT	1944 a	1949 c	1486 g.	1929 t	368 others
ORIGIN	/organism="unknown"				

```
Query Match          2.2%;      Score 35;      DB 22;      Length 7218;
Best Local Similarity 2.1%;      Pred. No. 9.13e-06;
Matches              3; Conservative 83; Mismatches 54; 7231-
```

[illegible]

b
1184 YYYYYYYYYYYYYYYYYY 1203
 : : : : :
1570 cataggacctctggcatttt 1589

RESULT	7	HUMFIBRLLN	9662 bp	mRNA	PRI
LOCUS		DEFINITION	Homo sapiens fibrillin mRNA, complete cds.		

ACCESSION	LI3923
IID	9306745
KEYWORDS	fibrillin.
SOURCE	human.
ORGANISM	Homo sapiens
REFERENCE	Eukaryotes: mitochondrial eukaryotes; Metazoa: Chordata; Vertebrata; Eutheria; Primates; Catarrhini; Homiidae; Homo. 1 (bases 1 to 9662)
AUTHORS	Pereira, L.V., D'Alessio, M., Ramirez, F., Lynch, J., Sykes, B., Pangillinan, T., and Bonadio, J.
TITLE	Genomic organization of the sequence coding for fibrillin, the defective gene product in Marfan syndrome
JOURNAL	Hum. Mol. Genet. 2, 961-968 (1993)
MEDLINE	93372860

FEATURES	Location/Qualifiers
source	1..9662
	/organism="Homo sapiens"
	/db_xref="taxon:9606"
5'UTR	1..133

5'UTR

5D

134. .8749

[illegible]

Query Match	2.2%;	Score 36;	DB 26;	Length 9662;
Best Local Similarity	66.4%;	Pred. No.	2.15e-06;	

```

Matches      13;  Conservative      0;  Mismatches      37;  Indels      0;  Gaps      0;

```

[illegible]

5' yucuccaadyalylactatggatagcccatagtcagccaccatgccaatgcttcaat 561

Db

4280 ACCATGGGATCTTACCGCTGTCTGTGCAGCCACCAACCATTGTTA...

CTTGGGCGTCCTGGCAAGGAATACACAGGTGATGG 4329

Qy 562 acccaagggtccttcaagtgtaaatgcaagcagggatataaaagccaatga 611

0
1
2
3
4
5
6
7
8
9

RESULTS


```

RESULT 11
LOCUS HSF1BUC 2156 bp RNA PRI 21-AUG-1995
DEFINITION H.sapiens mRNA for fibulin-1 C.
ACCESSION X53743
NID g31418
KEYWORDS fibulin-1 C; glycoprotein.
SOURCE human.
ORGANISM Homo sapiens
Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata;
Vertebrata; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 2156)
AUTHORS Argaves, W.S.
TITLE Direct Submission
JOURNAL Submitted (03-JUL-1990) Argaves W.S., American Red Cross, 15601
Crabbs Branch Way, Rockville, MD 20855, USA
REFERENCE 2 (bases 1 to 2156)
AUTHORS Argaves, W.S., Tran, H., Burgess, W.H. and Dickerson, K.
TITLE Fibulin is an extracellular matrix and plasma glycoprotein with
repeated domain structure
JOURNAL J. Cell Biol. 111 (6 Pt 2), 3155-3164 (1990)
MEDLINE 91100426
REFERENCE 3 (bases 1 to 2156)
AUTHORS Korenberg, J.R., Chen, X.N., Tran, H. and Argaves, W.S.
TITLE Localization of the human gene for fibulin-1 (FBLN1) to chromosome
band 22q13.3
JOURNAL Cytogenet. Cell Genet. 68 (3-4), 192-193 (1995)
MEDLINE 95145011
FEATURES
source
Location/Qualifiers
1..2156
/organism="Homo sapiens"
/db_xref="taxon:9606"
1..2156
/evidence="experimental"
11..86
/sig_peptide
CDS 11..2062
/codon_start=1
/product="fibulin-1 C"
/db_xref="PID:g31419"
/db_xref="SWISS-PROT:P23144"
/translation="MERAAPSRVPLPILLGLGALLAGVDVLEACCAADGHRMA
THQKDCSPYATESKCEKRVQCCCHSLEELHCATGTSILANEQDRCATPHGDNASLE
ATFKRCHCCLLGRAAQAGQSCSEYSLMVGYCGQGVFRACCVKSOETGDLVDVGLQE
DKIIEVEEEDPYLNDRCRGKQCKQCRDTGDEVVSCFVGYQLSDGVSCEDVN
ECITGSHCRIGESCINTVGSFRCDSSCGTGVELTDNSCKDIDECESGHNCILPD
FICONTLGSFRCPKLOCKSGFIODALNCIDINCLISAPCIPIGHTCINTEGTYTC
QKNVPCNGRGYHLNEEGTRCVDDCAPPAEPCGKGRVCSVPSGFRCECTGYFDG
ISRMVDVNECORYPRLCGHKCENTILGSCSVGFRLSDVGRSCEDINCCSSPC
SQECANVYGSQCYRRGYQLSDVDGVTCEIDECALPTGHHICSYRCINIPGSFQCS
CPSSGYRLAPNRCODIDECVTGTHNCININETCFNIQAFRLAFCEPENVRRAAT
RCEPLCHENRECSKLPRIYHLSPPTNIQAPAVFMGPPSSAVPDSQMLATGG
NEEGFFTRKVSHPHSGVVALPKVPEPRDLLLTVMDSLRHGTVSFVAKLFIIVSAE
L"
mat_peptide 87..2059
polyA_site 2156
BASE COUNT 420 a 654 c 647 g 435 t
ORIGIN
Query Match 2.1%; Score 33; DB 26; Length 2156;
Best Local Similarity 66.3%; Pred. No. 1.56e-04;
Matches 67; Conservative 0; Mismatches 34; Indels 0; Gaps 0;
Db 1526 CAGTGCAGCTGCCCTCGTGTGCTAGAGCTGGCCCCCAATGGCGCAACTGCCAAGAC 1585
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 310 cagtgcctgtgtccatctcaggactccgctggccccaaatggaagagactgtctagat 369
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1586 ATTGATGAGTGTGATGGCATCCACAACTGCTCCATCAA 1626
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 370 attgatgaatgtccctcgtgtaagtcattctgtccctacaa 410
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 12
LOCUS HSF1BUC 2349 bp RNA PRI 21-AUG-1995
DEFINITION H.sapiens mRNA for fibulin-1 A.
ACCESSION X53741
NID g31414
KEYWORDS fibulin-1 A; glycoprotein.
SOURCE human.
ORGANISM Homo sapiens
Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata;
Vertebrata; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 2349)
AUTHORS Argaves, W.S.
TITLE Direct Submission
JOURNAL Submitted (03-JUL-1990) Argaves W.S., American Red Cross, 15601
Crabbs Branch Way, Rockville, MD 20855, USA
REFERENCE 2 (bases 1 to 2349)
AUTHORS Argaves, W.S., Tran, H., Burgess, W.H. and Dickerson, K.
TITLE Fibulin is an extracellular matrix and plasma glycoprotein with
repeated domain structure
JOURNAL J. Cell Biol. 111 (6 Pt 2), 3155-3164 (1990)
MEDLINE 91100426
REFERENCE 3 (bases 1 to 2349)
AUTHORS Korenberg, J.R., Chen, X.N., Tran, H. and Argaves, W.S.
TITLE Localization of the human gene for fibulin-1 (FBLN1) to chromosome
band 22q13.3
JOURNAL Cytogenet. Cell Genet. 68 (3-4), 192-193 (1995)
MEDLINE 95145011
FEATURES
source
Location/Qualifiers
1..2349
/organism="Homo sapiens"
/db_xref="taxon:9606"
1..2349
/tissue_type="placenta"
/chromosome="22q13.3"
1..2349
/evidence="experimental"
11..85
/sig_peptide
CDS 11..1711
/codon_start=1
/product="fibulin-1 A"
/db_xref="PID:g31415"
/db_xref="SWISS-PROT:P23142"
/translation="MERAAPSRVPLPILLGLGALLAGVDVLEACCAADGHRMA
THQKDCSPYATESKCEKRVQCCCHSLEELHCATGTSILANEQDRCATPHGDNASLE
ATFKRCHCCLLGRAAQAGQSCSEYSLMVGYCGQGVFRACCVKSOETGDLVDVGLQE
DKIIEVEEEDPYLNDRCRGKQCKQCRDTGDEVVSCFVGYQLSDGVSCEDVN
ECITGSHCRIGESCINTVGSFRCDSSCGTGVELTDNSCKDIDECESGHNCILPD
FICONTLGSFRCPKLOCKSGFIODALNCIDINCLISAPCIPIGHTCINTEGTYTC
QKNVPCNGRGYHLNEEGTRCVDDCAPPAEPCGKGRVCSVPSGFRCECTGYFDG
ISRMVDVNECORYPRLCGHKCENTILGSCSVGFRLSDVGRSCEDINCCSSPC
SQECANVYGSQCYRRGYQLSDVDGVTCEIDECALPTGHHICSYRCINIPGSFQCS
CPSSGYRLAPNRCODIDECVTGTHNCININETCFNIQAFRLAFCEPENVRRAAT
RCEPLCHENRECSKLPRIYHLSPPTNIQAPAVFMGPPSSAVPDSQMLATGG
NEEGFFTRKVSHPHSGVVALPKVPEPRDLLLTVMDSLRHGTVSFVAKLFIIVSAE
L"
mat_peptide 86..1708
polyA_site 2349
BASE COUNT 520 a 668 c 682 g 479 t
ORIGIN
Query Match 2.1%; Score 33; DB 26; Length 2349;
Best Local Similarity 66.3%; Pred. No. 1.56e-04;
Matches 67; Conservative 0; Mismatches 34; Indels 0; Gaps 0;
Db 1526 CAGTGCAGCTGCCCTCGTGTGCTAGAGCTGGCCCCCAATGGCGCAACTGCCAAGAC 1585
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 310 cagtgcctgtgtccatctcaggactccgctggccccaaatggaagagactgtctagat 369
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1586 ATTGATGAGTGTGATGGCATCCACAACTGCTCCATCAA 1626
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 370 attgatgaatgtccctcgtgtaagtcattctgtccctacaa 410
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 13
LOCUS U01244 2359 bp mRNA PRI 18-OCT-1996

```


Qy	496	tgtatagatataaaatgaa	tgtactatgga	gccatacgtgcagccacatgccaatg	555
Db	4366	ANGAACCATCGGTCTATCGCTGTCTCTGTAAGGATGGCTATACAGGGGATCG	4421		
Qy	556	ttcaataccacaggtcttcaagtgtaaa	tgcaagcagggata	taaaagccaatgg	611

Search completed: Thu Apr 29 03:47:27 1999
Job time : 2515 secs.

Query Match	2.0%	Score 32;	DB 29;	Length 9830;
Best Local Similarity	63.8%	Pred. No.	6,25E-04;	
Matches	74;	Conservative	0;	Mismatches 42;
			Indels	0;
			Gaps	0;
Db	4306	TGCACAGATCTGGATCAATGCTCTATATGGNACCACCATGTGGCAGCCAAACACCGGACTGC	4365	


```

RESULT      2
ID   Q51746 standard; cDNA; 91 BP.
AC   Q51746;
DE   31-MAY-1994 (first entry)
DT   Oligonucleotide probe MK14-A
KW   Oligonucleotide; DNA probe; mycobacteria; disease diagnosis;
OS   ss.
PI   Synthetic.
PN   EP-571911-A.
PD   01-DEC-1993.
PF   24-MAY-1993; 108325.
PR   (BECT ) BECTON DICKINSON CO.
PA   Shank DD, Spears PA;
PI   WPI; 93-378844/48.
PT   New oligo:nucleotide probes specific for Mycobacteria - used for
    detection and amplification of Mycobacteria nucleic acid in
    samples
    Claim 3; Page 14; 23pp; English.
CC   Oligonucleotide probe MK14-A consists of nucleotides 5-95 of MK14
    (Q51735). It hybridized to all spp. of mycobacteria tested, but
    cross reacted to a few non-mycobacterial spp. The probe may
    be useful as an initial screen for mycobacterial infection.
    See also Q51735-45 and Q51747-59.
SQ   Sequence 91 BP; 5 A; 17 C; 15 G; 4 T;

Query Match          2.6%; Score 41; DB 9; Length 91;
Best Local Similarity 5.6%; Pred. NO. 2.86e-10;
Matches 3; Conservative 47; Mismatches 3; Indels 1; Gaps 1;

Db 10 gssvhyvvhvshhsbhvhvhhsvvvvhhvhvhh-vhyhyvvsvtc 62
    |:::| ::::| ::::| ::::| ::::| ::::| ::::| ::::| ::::| ::::|
Cp 981 gcgaattacactgcttcattcccttagggaaacacatctctcgaggt 928

RESULT      3
ID   N81164 standard; DNA; 204 BP.
AC   N81164;
DE   08-NOV-1990 (first entry)
DT   Base substituted E.coli beta-galactosidase alpha-fragment.
KW   E.coli beta galactosidase alpha-fragment; base substitutions; ss.
OS   Escherichia coli.
Key Location/Qualifiers
misc_feature 19..69 /tag= a
primer_bind 187..204 /function=multiple cloning site
                    /*tag= b
    EP-285123-A.
    PN
    PD 05-MAY-1988.
    PF 30-MAR-1988; 105163.
    PR 03-APR-1987; US-034819.
    PA (SUSO) SUOMEN SOKERI OY.
    PI Lehtovaara P, Knowles J, Koivula A, Bamford J, Reinikainen T;
    PT WPI; 88-279927/40.
    PT Introducing random point mutations into nucleic acids -
    by prepn of single stranded template, annealing a primer, elongation,
    PT misincorporation, completion of molecules and screening.
    PS Disclosure; p; English.
    CC Random point mutations were introduced into the alpha fragment of
    CC E.coli beta-galactosidase. The wild type sequence was obtained as a
    CC single stranded template and an oligonucleotide was hybridised to
    CC it to generate a popn of DNA molecules which terminate at all
    CC possible nucleotide positions within a specified region. The
    CC variable 3' ends generated in this way are used as primers for
    CC reverse transcriptase. Nucleotides are misincorporated by the
    CC transcriptase and the molecules are completed to forms that can be
    CC amplified and then expressed in a suitable host-vector system.
    CC The sequence covers all 176 diffit base substitutions, most of which
    CC occurred singularly in any given mutant.
    CC See also P80575.

```

```

KW direct; rapid; detection; screening; treatment; generic; ss.
OS Synthetic.
FH Key Location/Qualifiers
FT misc_feature 55..60
FT /*tag= a
FT /*note= "this sequence represents 'Z'; Z can be a
FT sequence of 6, 9 or 12 nucleotides (see
FT comments)"
PN WO9418318-A.
PD 18-AUG-1994.
PF 01-FEB-1994. U00977.
PR 01-FEB-1993; US-013416.
PR 30-DEC-1993; US-176500.
PR 31-JAN-1994; US-189331.
PA (UYN-) UNIV NORTH CAROLINA.
PI Fowlkes DM, Kay BK;
DR WPI; 94-279739/34.
DR P-PSDB; R65153.
PT Identifying proteins or peptide(s) which bind a ligand - by
PT screening a recombinant vector library expressing fusion proteins
PT comprising a binding domain and an effector domain
PS Disclosure; Page 35; 255pp; English.
CC Synthetic Affinity Reagents) peptides. This generic formula can also be
CC represented as follows: X(NNB)16(TGC)(NNB)12(NNB)16(TGC)(NNB)10Y..X
CC and Y are flanking restriction sites (X is not the same as Y) that are
CC not specified further. Other generic sequences are shown in Q70466-68.
CC Other specific peptides generated by these generic sequences are shown in
CC R65151-54. TSARs are concatenated heterofunctional proteins or peptides,
CC comprising at least two functional regions - a binding domain with
CC affinity for a ligand and a second effector peptide portion that is
CC chemically or biologically active. They may further comprise a linker
CC peptide between the 2 domains. The oligonucleotides are also designed so
CC that the expressed peptide contains 2 or 4 cysteine residues positioned
CC in, or flanking, the unpredicted or variant residues. These residues
CC confer some degree of conformational rigidity to the peptides. The TSARs
CC or compsns. comprising a TSAR binding domain can be used in vivo to
CC deliver a chemically or biologically active moiety, eg. metal ion,
CC radioisotope, peptide, toxin or enzyme, to the specific target or on the
CC cell. They can also replace the function of macromolecules, eg.
CC monoclonal or polyclonal antibodies and therefore circumvent the need
CC for complex methods of hybridoma formation or in vivo antibody
CC production. The TSARs are easily characterised and have designed activity
CC allowing direct and rapid detection in a screening process.
SQ Sequence 114 BP; 0 A; 2 C; 2 G; 2 T;

Query Match 2.1%; Score 34; DB 12; Length 114;
Best Local Similarity 1.9%; Pred. No. 3.70e-06;
Matches 2; Conservative 33; Mismatches 71; Indels 0; Gaps 0;

Db 3 bnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnnnn 62
Cp 940 cctcgcaggcttcgtctctatgctcattcttcagggtttctctctctctctcat 881
Db 63 bnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnnn 108
Cp 880 cctcagccctctctctctctctctctctctctctctctctctctctctctcc 835

RESULT 6
ID Q70468 standard; DNA; 114 BP.
AC Q70468;
DE 05-APR-1995 (first entry)
KW Generic DNA sequence to generate a random TSAR peptide library.
KW TSAR; totally synthetic affinity reagent; synthetic; binding domain;
KW effector domain; concatenated heterofunctional protein; linker;
KW direct; rapid; detection; screening; treatment; generic; ss.
OS Synthetic.
FH Key Location/Qualifiers
FT misc_feature 55..60
FT /*tag= a
FT /*note= "this sequence represents 'Z'; Z can be a
FT sequence of 6, 9 or 12 nucleotides (see
FT comments)"
PN WO9418318-A.
PD 18-AUG-1994.
PF 01-FEB-1994. U00977.
PR 01-FEB-1993; US-013416.
PR 30-DEC-1993; US-176500.
PR 31-JAN-1994; US-189331.
PA (UYN-) UNIV NORTH CAROLINA.
PI Fowlkes DM, Kay BK;
DR WPI; 94-279739/34.
DR P-PSDB; R65153.
PT Identifying proteins or peptide(s) which bind a ligand - by
PT screening a recombinant vector library expressing fusion proteins
PT comprising a binding domain and an effector domain
PS Disclosure; Page 35; 255pp; English.
CC Synthetic Affinity Reagents) peptides. This generic formula can also be
CC represented as follows: X(NNB)16(TGC)(NNB)12(NNB)16(TGC)(NNB)10Y..X
CC and Y are flanking restriction sites (X is not the same as Y) that are
CC not specified further. Other generic sequences are shown in Q70466-68.
CC Other specific peptides generated by these generic sequences are shown in
CC R65151-54. TSARs are concatenated heterofunctional proteins or peptides,
CC comprising at least two functional regions - a binding domain with
CC affinity for a ligand and a second effector peptide portion that is
CC chemically or biologically active. They may further comprise a linker
CC peptide between the 2 domains. The oligonucleotides are also designed so
CC that the expressed peptide contains 2 or 4 cysteine residues positioned
CC in, or flanking, the unpredicted or variant residues. These residues
CC confer some degree of conformational rigidity to the peptides. The TSARs
CC or compsns. comprising a TSAR binding domain can be used in vivo to
CC deliver a chemically or biologically active moiety, eg. metal ion,
CC radioisotope, peptide, toxin or enzyme, to the specific target or on the
CC cell. They can also replace the function of macromolecules, eg.
CC monoclonal or polyclonal antibodies and therefore circumvent the need for
CC complex methods of hybridoma formation or in vivo antibody production.
CC The TSARs are easily characterised and have designed activity allowing
CC direct and rapid detection in a screening process.
SQ Sequence 114 BP; 0 A; 2 C; 2 G; 2 T;

Query Match 2.1%; Score 34; DB 12; Length 114;
Best Local Similarity 1.9%; Pred. No. 3.70e-06;
Matches 2; Conservative 33; Mismatches 71; Indels 0; Gaps 0;

Db 3 bnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnnnn 62
Cp 940 cctcgcaggcttcgtctctatgctcattcttcagggtttctctctctctctcat 881
Db 63 bnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnnn 108
Cp 880 cctcagccctctctctctctctctctctctctctctctctctctctctctcc 835

RESULT 7
ID Q70465 standard; DNA; 114 BP.
AC Q70465;
DE 05-APR-1995 (first entry)
KW Generic DNA sequence to generate a random TSAR peptide library.
KW TSAR; totally synthetic affinity reagent; synthetic; binding domain;
KW effector domain; concatenated heterofunctional protein; linker;
KW direct; rapid; detection; screening; treatment; generic; ss.
OS Synthetic.
FH Key Location/Qualifiers
FT misc_feature 55..60
FT /*tag= a
FT /*note= "this sequence represents 'Z'; Z can be a
FT sequence of 6, 9 or 12 nucleotides (see
FT comments)"
PN WO9418318-A.
PD 18-AUG-1994.
PF 01-FEB-1994. U00977.
PR 01-FEB-1993; US-013416.
PR 30-DEC-1993; US-176500.
PR 31-JAN-1994; US-189331.

```

(UYN-) UNIV NORTH CAROLINA.
 FI Fowles DM, Kay BK;
 DR WPI; 94-279739/34.
 DR P-PSDB; R65150 and R65151.
 PT Identifying proteins or peptide(s) which bind a ligand - by
 PT screening a recombinant vector library expressing fusion proteins
 PT comprising a binding domain and an effector domain
 PS Disclosure; Page 35; 255pp; English.
 CC Q70465 is a generic DNA sequence used to generate random TSAR (Totally
 CC Synthetic Affinity Reagents) peptides. This generic formula can also be
 CC represented as follows: X(NNB)6(TGC)(NNB)12(TGC)(NNB)3Y. X
 CC and Y are flanking restriction sites (X is not the same as Y) that are
 CC not specified further. Other generic sequences are shown in Q70466-68.
 CC Other specific peptides generated by these generic sequences are shown in
 CC R65151-54. TSARs are concatenated heterofunctional proteins or peptides,
 CC comprising at least two functional regions - a binding domain with
 CC affinity for a ligand and a second effector peptide portion that is
 CC chemically or biologically active. They may further comprise a linker
 CC peptide between the 2 domains. The oligonucleotides are also designed so
 CC that the expressed peptide contains 2 or 4 cysteine residues positioned
 CC in, or flanking, the unpredicted or variant residues. These residues
 CC confer some degree of conformational rigidity to the peptides. The TSARs
 CC or compsns. comprising a TSAR binding domain can be used in vivo to
 CC deliver a chemically or biologically active moiety, eg. metal ion,
 CC radioisotope, peptide, toxin or enzyme, to the specific target or on the
 CC cell. They can also replace the function of macromolecules, eg.
 CC monoclonal or polyclonal antibodies and therefore circumvent the need
 CC for complex methods of hybridoma formation or in vivo antibody
 CC production. The TSARs are easily characterized and have designed
 CC activity allowing direct and rapid detection in a screening process.
 SQ Sequence 114 BP; 0 A; 2 C; 2 G; 2 T;

Query Match 2.1%; Score 34; DB 12; Length 114;
 Best Local Similarity 3.6%; Pred. No. 3.70e-06;
 Matches 4; Conservative 33; Mismatches 75; Indels 0; Gaps 0;

Db 3 bnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnnn 62
 Cp 763 tctgtggtggtctggggtacatttttgcctctcttctgtgtgtttgt 704
 Db 63 bnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnn 114
 Cp 703 gagcaagcaactctgtattctgtctgttgatgtgtaccagggtctgaggac 652

RESULT 8
 ID Q70469 standard; DNA; 114 BP.
 AC Q70469;
 DE 07-APR-1995 (first entry)
 KW TSAR; totally synthetic affinity reagent; synthetic; binding domain;
 KW effector domain; concatenated heterofunctional protein; linker;
 KW direct; rapid; detection; screening; treatment; generic; ss.
 OS Synthetic.
 FH Key Location/Qualifiers
 FT mat_feature 55..60
 FT misc_feature
 FT /tag= a
 FT /note= "this sequence represents '2'; 2 can be a
 FT sequence of 6,9 or 12 nucleotides (see
 FT comments)"
 FT
 PN W09418318-A.
 PD 18-AUG-1994.
 PF 01-FEB-1994; U00977.
 PR 01-FEB-1993; US-013416.
 PR 30-DEC-1993; US-176500.
 PR 31-JAN-1994; US-189331.
 PA (UYN-) UNIV NORTH CAROLINA.
 PI Fowles DM, Kay BK.
 DR WPI; 94-279739/34.
 PT Identifying proteins or peptide(s) which bind a ligand - by
 PT screening a recombinant vector library expressing fusion proteins
 PT comprising a binding domain and an effector domain
 PS Disclosure; Page 35; 255pp; English.

Q70469 is a generic DNA sequence used to generate random TSAR peptide
 This generic formula can be represented as follows: X(TGC)(NNB)10-
 (TGC)(NNB)6Z(NNB)2(TGC)(NNB)14(TGC)Y. X and Y are flanking restriction
 sites (X is not the same as Y) that are not specified further. This
 sequence generates peptides that are cloverleaf in structure. Other
 generic sequences are shown in Q70465-68. Other specific peptides
 generated by these generic sequences are shown in R65150-54. TSARs are
 concatenated heterofunctional proteins or peptides, comprising at least
 two functional regions - a binding domain with affinity for a ligand and
 a second effector peptide portion that is chemically or biologically
 active. They may further comprise a linker peptide between the 2 domains.
 The oligonucleotides are also designed so that the expressed peptide
 contains 2 or 4 cysteine residues positioned in, or flanking, the
 unpredicted or variant residues. These residues confer some degree of
 conformational rigidity to the peptides. The TSARs or compsns. comprising
 a TSAR binding domain can be used in vivo to deliver a chemically or
 biologically active moiety, eg. metal ion, radioisotope, peptide, toxin
 or enzyme, to the specific target or on the cell. They can also replace
 the function of macromolecules, eg. monoclonal or polyclonal antibodies
 and therefore circumvent the need for complex methods of hybridoma
 formation or in vivo antibody production. The TSARs are easily
 characterized and have designed activity allowing direct and rapid
 detection in a screening process.
 SQ Sequence 114 BP; 0 A; 4 C; 4 G; 4 T;

Query Match 2.1%; Score 33; DB 12; Length 114;
 Best Local Similarity 5.4%; Pred. No. 1.36e-05;
 Matches 6; Conservative 31; Mismatches 74; Indels 0; Gaps 0;

Db 2 gcnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnnn 61
 Cp 1004 gtttccttggaccaggaatcaggccggaattcacctgttcattcactagggaacac 945
 Db 62 nbnnbgtcnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnn 112
 Cp 944 acatctctcaggtctgcctctctatctctatctctcaggtctctctctctctct 894

RESULT 9
 ID Q11010 standard; cDNA; 2200 BP.
 AC Q11010;
 DE 21-MAY-1991 (first entry)
 DE Fibulin C.
 KW Beta-1 integrin; adhesion; receptor; fibronectin; ss.
 OS Homo sapiens.
 FH Key Location/Qualifiers
 FT signal_peptide 11..97
 FT mat_peptide /tag= a
 FT /tag= b
 FT /product= fibulin C
 FT misc_feature 1708
 FT /tag= c
 FT poly_a_signal /label= point of divergence
 FT 2126..2131
 FT /tag= d
 PN W09102755-A.
 PD 07-MAR-1991.
 PF 17-AUG-1990; U04662.
 PR 18-AUG-1989; US-395773.
 PR (LJOL-) LA JOLLA CANCER RES.
 PA (AMNA-) AMER NAT RED CROSS.
 PI Ruoslahti EI, Argraves WS;
 DR WPI; 91-087250/12.
 DR N-PSDB; R11150.
 PT Purified fibulin, DNA encoding it and antibodies reactive with it
 PT - useful as diagnostic and therapeutic component.
 PS Claim 10; Fig 3; 56pp; English.
 CC The fibulin C cDNA was sequenced from a clone isolated from lambda
 CC gtl human placental cDNA library using a polyclonal antiserum.
 CC The three forms (A, B and C) are identical from their 5' ends to a
 CC divergence point at posn. 1707, after which they are distinct
 CC through to the poly(A) tail. The cDNA can be used to express


```

FT misc_feature 55..60
FT /*tag= a
FT /note= "this sequence represents 'Z'; Z can be a
FT sequence of 6, 9 or 12 nucleotides (see
FT comments)"
PN WO9418318-A.
PD 18-AUG-1994.
PF 01-FEB-1994; U00977.
PR 01-FEB-1993; US-013416.
PR 30-DEC-1993; US-176500.
PR 31-JAN-1994; US-189331.
PA (UYNC-) UNIV NORTH CAROLINA.
PI Fowlkes DM, Kay BK;
DR WPI: 94-279739/34.
DR P-PSDB; R65150 and R65151.
PT Identifying proteins or peptide(s) which bind a ligand - by
PT screening a recombinant vector library expressing fusion proteins
PT comprising a binding domain and an effector domain
PS Disclosure; Page 35; 255pp; English.
CC Synthetic Affinity Reagents) peptides. This generic formula can also be
CC represented as follows: X(NNB)6(TGC)(NNB)12(TGC)(NNB)3Y. X
CC and Y are flanking restriction sites (X is not the same as Y) that are
CC not specified further. Other generic sequences are shown in
CC Q70465.
CC R65151-54. TSARs are concatenated heterofunctional proteins or peptides,
CC comprising at least two functional regions - a binding domain with
CC affinity for a ligand and a second effector peptide portion that is
CC chemically or biologically active. They may further comprise a linker
CC that the expressed peptide contains 2 or 4 cysteine residues positioned
CC in, or flanking, the unpredicted or variant residues. These residues
CC confer some degree of conformational rigidity to the peptides. The TSARs
CC deliver a chemically or biologically active moiety, eg. metal ion,
CC cell. They can also replace the function of macromolecules, eg.
CC monoclonal or polyclonal antibodies and therefore circumvent the need
CC for complex methods of hybridoma formation or in vivo antibody
CC production. The TSARs are easily characterised and have designed
CC activity allowing direct and rapid detection in a screening process.
SQ Sequence 114 BP; 0 A; 2 C; 2 G; 2 T;

Query Match 2.0%; Score 32; DB 12; Length 114;
Best Local Similarity 3.6%; Pred. No. 4.97e-05;
Matches 4; Conservative 32; Mismatches 76; Indels 0; Gaps 0;

Db 3 bnnbnnbnnbnnbgtcnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnnn 62
1038 taticgttgactgcagcttcaatcgtggtctgtgactggaacagcagatagaaga 1097

Db 63 bnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnb 114
1098 tgatttgactgaatccctgcgtgatcgatagataatgctattgcttatatg 1149

RESULT 13
ID Q70466 standard; DNA; 114 BP.
AC Q70466;
DT 05-APR-1995 (first entry)
DE Generic DNA sequence to generate a random TSAR-9 peptide library.
KW TSAR; totally synthetic affinity reagent; synthetic; binding domain;
KW effector domain; concatenated heterofunctional protein; linker;
KW direct; rapid; detection; screening; treatment; generic; ss.
OS Synthetic.
FH Key Location/Qualifiers
FT misc_feature 55..60
FT /*tag= a
FT /note= "this sequence represents 'Z'; Z can be a
FT sequence of 6, 9 or 12 nucleotides (see
FT comments)"
PN WO9418318-A.
PD 18-AUG-1994.

```

```

PF 01-FEB-1994; U00977.
PR 01-FEB-1993; US-013416.
PR 30-DEC-1993; US-176500.
PR 31-JAN-1994; US-189331.
PA (UYNC-) UNIV NORTH CAROLINA.
PI Fowlkes DM, Kay BK;
DR WPI: 94-279739/34.
DR P-PSDB; R65152.
PT Identifying proteins or peptide(s) which bind a ligand - by
PT screening a recombinant vector library expressing fusion proteins
PT comprising a binding domain and an effector domain
PS Disclosure; Page 35; 255pp; English.
CC Q70466 is a generic DNA sequence used to generate random TSAR (Totally
CC Synthetic Affinity Reagents) peptides. This generic formula can also be
CC represented as follows: X(NNB)1(TGC)(NNB)10(TGC)2(NNB)42(NNB)8(TGC)(NNB)
CC -9Y. X and Y are flanking restriction sites (X is not the same as Y)
CC that are not specified further. Other generic sequences are shown in
CC Q70466-68. Other specific peptides generated by these generic sequences
CC are shown in R65151-54. TSARs are concatenated heterofunctional proteins
CC or peptides, comprising at least two functional regions - a binding
CC domain with affinity for a ligand and a second effector peptide portion
CC that is chemically or biologically active. They may further comprise a
CC linker peptide between the 2 domains. The oligonucleotides are also
CC designed so that the expressed peptide contains 2 or 4 cysteine residues
CC positioned in, or flanking, the unpredicted or variant residues. These
CC residues confer some degree of conformational rigidity to the peptides.
CC The TSARs or compsns. comprising a TSAR binding domain can be used in
CC vivo to deliver a chemically or biologically active moiety, eg. metal
CC ion, radioisotope, peptide, toxin or enzyme, to the specific target or
CC on the cell. They can also replace the function of macromolecules, eg.
CC monoclonal or polyclonal antibodies and therefore circumvent the need for
CC complex methods of hybridoma formation or in vivo antibody production.
CC The TSARs are easily characterised and have designed activity allowing
CC direct and rapid detection in a screening process.
SQ Sequence 114 BP; 0 A; 4 C; 4 G; 4 T;

Query Match 2.0%; Score 32; DB 12; Length 114;
Best Local Similarity 5.7%; Pred. No. 4.97e-05;
Matches 6; Conservative 30; Mismatches 70; Indels 0; Gaps 0;

Db 9 bnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnb 68
997 ttggaccagaatcaggccgaatcaccttccttcattcaccttagggaaacacatctc 938

Db 69 bnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnb 114
937 ctgcaggcttcgctctctctctctctctctctcagggcttctcttc 892

RESULT 14
ID Q70470 standard; DNA; 114 BP.
AC Q70470;
DT 10-APR-1995 (first entry)
DE Generic DNA sequence to generate a random TSAR peptide library.
KW TSAR; totally synthetic affinity reagent; synthetic; binding domain;
KW effector domain; concatenated heterofunctional protein; linker;
KW direct; rapid; detection; screening; treatment; generic; ss.
OS Synthetic.
FH Key Location/Qualifiers
FT misc_feature 55..60
FT /*tag= a
FT /note= "encoded by Z (see comments)"
PN WO9418318-A.
PD 18-AUG-1994.
PF 01-FEB-1994; U00977.
PR 01-FEB-1993; US-013416.
PR 30-DEC-1993; US-176500.
PR 31-JAN-1994; US-189331.
PA (UYNC-) UNIV NORTH CAROLINA.
PI Fowlkes DM, Kay BK;
DR WPI: 94-279739/34.
DR P-PSDB; R58378.
PT Identifying proteins or peptide(s) which bind a ligand - by

```

CC affinity for a ligand and a second effector peptide portion that is
CC chemically or biologically active. They may further comprise a linker
CC peptide between the 2 domains. The TSARs or compans. comprising a TSAR
CC binding domain can be used in vivo to deliver a chemically or
CC biologically active moiety, eg. metal ion, radioisotope, peptide, toxin
CC or enzyme, to the specific target or on the cell. They can also replace
CC the function of macromolecules, eg. monoclonal or polyclonal antibodies
CC and therefore circumvent the need for complex methods of hybridoma
CC formation or in vivo antibody production. The TSARs are easily
CC characterised and have designed activity allowing direct and rapid
CC detection in a screening process.

SQ Sequence 114 BP; 8 A; 16 C; 0 G; 0 T;

Query Match 1.9%; Score 30; DB 12; Length 114;
Best Local Similarity 12.7%; Pred. No. 6.31e-04;
Matches 13; Conservative 25; Mismatches 64; Indels 0; Gaps 0

Dbb 1 cacacnbnnnn 60
||| : : : : : : : : : : : : : : : : : :
Cp 945 cacatctcctgcaggcttcgctctctatgtcatcttccaggctttctctctttt 886
||| : : : : : : : : : : : : : : : : : :

Dbb 61 cacacnbn 102
||| : : : : : : : : : : : : : : : : : :

Cp 885 ctcatctctcaagcccctttctattctctctctctctctctctccctt 844
||| : : : : : : : : : : : : : : : : : :

Search completed: Sat Apr 24 23:20:00 1999

Job time : 273 secs.

WIPSLH
(TM)

Release 3.1A John F. Collins, BioComputing Research Unit.
Copyright (c) 1993-1998 University of Edinburgh, U.K.
Distribution rights by Oxford Molecular Ltd

srch_nn n.a. - n.a. database search, using Smith-Waterman algorithm
Run on: Sat Apr 24 22:35:35 1999; MasPar time 2350.24 Seconds
Tabular output not generated. 1227.509 Million cell updates/sec

Title: >US-08-968-800-2
Description: (1-1611) from US08968800.seq
Perfect Score: 1607
N.A. Sequence: 1 gctggagaagaacagca.....aattactagctgaaaaatg 1611
Comp: ccgaccttttttctgtt.....ttaagatgacatttttaac

Scoring table:
Gap 6
TABLE default

Nmatch STD : Dbase 0; Query 0

Searched: 2275026 seqs, 895388244 bases x 2

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database: emb1-est56
Database: genbank-est109
1:em_est1 2:em_gss1 3:em_gss2 4:em_gss3
5:gb_est1 6:gb_est10 7:gb_est11 8:gb_est12 9:gb_est13
10:gb_est14 11:gb_est15 12:gb_est16 13:gb_est17
14:gb_est18 15:gb_est19 16:gb_est2 17:gb_est20
18:gb_est21 19:gb_est3 20:gb_est4 21:gb_est5 22:gb_est6
23:gb_est7 24:gb_est8 25:gb_est9 26:gb_gss1 27:gb_gss2
28:gb_gss3 29:gb_gss4

Statistics: Mean 11.687; Variance 2.100; scale 5.565

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	ID	Description	Pred. No.
1	555	34.5	626	AA195615	zr37b10.r1 Soares NHM	0.00e+00
2	451	28.1	508	AA025649	ze85d03.r1 Soares feta	0.00e+00
3	438	27.3	472	AA436507	zv08c07.r1 Soares NHM	0.00e+00
4	398	24.8	446	16 R99817	yq69b11.r1 Homo sapien	0.00e+00
5	355	22.1	473	16 R76984	yf66d08.r1 Homo sapien	0.00e+00
6	353	22.0	411	16 R64051	y122c04.r1 Homo sapien	0.00e+00
7	334	22.0	454	16 H03936	y744a05.r1 Homo sapien	0.00e+00
8	332	20.7	429	22 AA009900	ze83d06.r1 Soares feta	0.00e+00
9	331	20.6	341	16 R68291	y106a03.r1 Homo sapien	0.00e+00
10	311	19.4	442	5 T65608	yc76b01.r1 Homo sapien	0.00e+00
11	286	18.4	406	18 AI057064	oz07b04.r1 Soares_feta	0.00e+00
12	285	17.7	344	21 W80521	zd90b04.r1 Soares_feta	0.00e+00
13	277	17.2	347	5 R11377	yf42e08.r1 Homo sapien	0.00e+00

14	252	15.7	374	5 R27678	yh64e08.r1 Homo sapien	0.00e+00
15	249	15.5	364	5 R27153	yh46e02.r1 Homo sapien	0.00e+00
16	247	15.4	256	19 HSC32G011	H. sapiens partial CDN	0.00e+00
17	248	15.4	325	5 R24466	yh48c12.r1 Homo sapien	0.00e+00
18	236	14.7	379	5 R27258	yh53d04.r1 Homo sapien	0.00e+00
19	234	14.6	424	5 R24180	yh29g10.r1 Homo sapien	0.00e+00
20	218	13.6	270	5 R32182	yh67f04.r1 Homo sapien	0.00e+00
21	218	13.6	388	16 R64033	y119h01.r1 Homo sapien	0.00e+00
22	197	12.3	215	6 C17012	Human placenta cDNA 5'	0.00e+00
23	186	11.6	353	5 R34484	yh85f01.r1 Homo sapien	0.00e+00
24	178	11.1	213	19 N39776	yh65a12.r1 Homo sapien	0.00e+00
25	137	8.5	527	11 AA707654	z129e09.sl Soares_feta	9.47e-230
26	118	7.3	420	18 AA913032	o134a06.sl Soares_NFL	5.66e-188
27	114	7.1	506	18 AI057014	oz06d10.sl Soares_feta	2.90e-179
28	96	6.0	474	17 AI023885	ow69c06.sl Soares_feta	1.58e-140
29	86	5.4	485	22 AA009486	ze83d06.sl Soares_feta	2.06e-119
30	77	4.8	471	21 AA025548	ze85d03.sl Soares_feta	1.00e-100
31	64	4.0	252	12 AA754459	97SN1787 Rice Immature	1.96e-74
32	32	3.8	232	12 AA754459	97SN1787 Rice Immature	1.64e-68
33	61	3.8	457	23 AA259930	vab3c02.r1 Soares_mous	1.64e-68
34	60	3.7	439	17 AI022137	ow68e06.sl Soares_feta	1.49e-66
35	43	2.7	247	12 AA754458	97SN1784 Rice Immature	1.09e-34
36	44	2.7	247	12 AA754458	97SN1784 Rice Immature	1.09e-36
37	44	2.7	423	11 AA702522	zi80d04.sl Soares_feta	1.80e-36
38	39	2.4	334	10 AA634649	np48a02.sl NCI_CGAP.Br	1.01e-27
39	35	2.2	2275	11 AF034173	Homo sapiens ntcon2 co	4.84e-21
40	33	2.1	360	6 C18091	Human placenta cDNA 5'	7.85e-18
41	33	2.1	408	16 R69676	y140f01.r1 Homo sapien	7.85e-18
42	33	2.1	420	16 R71008	y150g07.r1 Homo sapien	7.85e-18
43	34	2.1	452	6 AA130066	z133c11.r1 Soares_preg	2.00e-19
44	34	2.1	467	16 R66735	y133b06.r1 Homo sapien	7.85e-18
45	33	2.1	479	16 H01614	yj21d02.r1 Homo sapien	7.85e-18

ALIGNMENTS

1
RESULT AA195615 626 bp mRNA EST 19-MAY-1997
LOCUS zr37b10.r1 Soares NHMPU S1 Homo sapiens cDNA clone 665563 5', mRNA
DEFINITION sequence.
ACCESSION AA195615
NID g1783792
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata.
Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Homnidae;
Homo.
REFERENCE 1 (bases 1 to 626)
AUTHORS Hillier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M., Holman,M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marra,M., Parsons,J., Rifkin,L., Rohlfing,T., Tan,F., Trevaskis,E., Waterston,R., Williamson,A., Wohlmann,P. and Wilson,R.
WashU-Merck EST Project
Unpublished (1995)
Contact: Wilson RK
WashU-Merck EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@wustl.wustl.edu
This clone is available royalty-free through LLNL; contact the IMAGE Consortium (info@image.llnl.gov) for further information.
Insert Length: 1094 Std Error: 0.00
High quality sequence stop: 477.
Location/Qualifiers
1. 626
/organism="Homo sapiens"
/note="Organ: mixed (see below); Vector: p7T3b-Pac (Pharmacia) with a modified polylinker; Site_1: Not I; Site_2: Eco RI; Equal amounts of plasmid DNA from three

FEATURES

source

[illegible]

CDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified pYT3 vector. Library went through one round of normalization. Library constructed by Bento Soares and M.Fatima Bonaldo.

ORGANISM

Homo sapiens
Eukaryotae; Metazoa; Eumetazoa; Bilateria; Coelomata; Deuterostomia; Chordata; Vertebrata; Gnathostomata; Osteichthyes; Sarcopterygii; Chonata; Tetrapoda; Amniota; Mammalia; Theria; Eutheria; Archonta; Primates; Catarrhini; Hominoidea; Homo.

REFERENCE

1 (bases 1 to 411)
Hillier, L., Clark, N., Dubuque, T., Elliston, K., Hawkins, M., Holman, M., Hultman, M., Kucaba, T., Le, M., Lennon, G., Marra, M., Parsons, J., Rifkin, L., Rohlfing, T., Soares, M., Tan, F., Trevasakis, E., Waterston, R., Williamson, A., Wohlmann, P. and Wilson, R.

TITLE

The WashU-Merck EST Project

JOURNAL

Unpublished (1995)

COMMENT

Contact: Wilson RK
WashU-Merck EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
High quality sequence stops: 314
Source: IMAGE Consortium, LLNL
This clone is available royalty-free through LLNL; contact the IMAGE Consortium (info@image.llnl.gov) for further information.

FEATURES

source
1..411
Location/Qualifiers
/organism="Homo sapiens"
/clone="139974"

BASE COUNT 121 a 80 c 107 g 100 t 3 others

ORIGIN

Query Match 22.0%; Score 353; DB 16; Length 411;
Best Local Similarity 95.8%; Pred. No. 0.00e+00;
Matches 392; Conservative 0; Mismatches 12; Indels 5; Gaps 5;

Db 2 CAAGGAAGCGCTCACTCCAACTGGAACATGAAGATTAATATCTCGGTGACTGC 61
QY 994 caaaggaaagcgcctaactccaaactggaacacgagatgagaaagatttgaatctcgttgactgc 1053
|||

Db 62 AGCTTCATCATGGATCTGTGACATGGAAACAGGATAGAGACATGATTTGACTGGAAT 121
QY 1054 agcttcaatcatggatctgtgactggaacacgagatgagaaagatttgaatctcgttgactgc 1113
|||

Db 122 CTGTGATCGAGATAATGCTATTGGCTTCTATATGGCAGTTCGGCTTGGCAGGTCCAC 181
QY 1114 cctgctgacgagataaagtctattgcttctatatgagcagttccgctgagcagtcac 1173
|||

Db 182 AAGAAGACATTTGGCGGATTTGAACTTCTCTACCTGACCTGCAACCCCAAGCAACTTC 241
QY 1174 atgaagacattggcggatgaaactctcctacctgacctgcaaccccaagcaacttc 1233
|||

Db 242 TGTTCCTCTTTGATTACGGCTGGCGGAGGACAAAGTCGGGAAACTTCGAGTGTGT 301
QY 1234 tgttgcctcttgattacggctggcggagacaaagtcggg-aaacttcgagttgtt 1292
|||

Db 302 GAAACACAGTACATGCTCGTGNATGGGAGAACACACGAGTGAGGTTGAAAGCTGGG 361
QY 1293 gaaacacagtaacaatgcccctggcatggg-agaagaccacgagtgaggaatgaaagtcg- 1350
|||

Db 362 AAGNAGGAGGAAATTCAGTCTTTTCAAGGACCTGTTTACCCAAAG 410
QY 1351 aagacagggaaaatt-cagttgtat-caaggaaactgacttaccacaaaag 1397
|||

RESULT

7 H03936 454 bp mRNA EST 20-JUN-1995

LOCUS

DEFINITION Y144a05.s1 Homo sapiens CDNA clone 151568 3'

ACCESSION

H03936

9866869

EST.

human clone=151568 library=Soares placenta NB2HP vector=pYT3D (Pharmacia) with a modified polylinker host=DH10B (ampicillin resistant) primer=Promega -2lm3 Rsite1-Not I Rsite2=Eco RI Female placenta obtained at birth (full term). 1st strand cDNA was primed with a Not I - Oligo(dT) primer [5' AACGGAAGATTCGCGCGCAGGAATTTTTTTTTTTTTTTT 3'], double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified pYT3 vector. Library went through one round of normalization. Library constructed by Bento Soares and M.Fatima Bonaldo.

ORGANISM

Homo sapiens
Eukaryotae; Metazoa; Eumetazoa; Bilateria; Coelomata; Deuterostomia; Chordata; Vertebrata; Gnathostomata; Osteichthyes; Sarcopterygii; Chonata; Tetrapoda; Amniota; Mammalia; Theria; Eutheria; Archonta; Primates; Catarrhini; Hominoidea; Homo.

REFERENCE

1 (bases 1 to 454)
Hillier, L., Clark, N., Dubuque, T., Elliston, K., Hawkins, M., Holman, M., Hultman, M., Kucaba, T., Le, M., Lennon, G., Marra, M., Parsons, J., Rifkin, L., Rohlfing, T., Soares, M., Tan, F., Trevasakis, E., Waterston, R., Williamson, A., Wohlmann, P. and Wilson, R.

TITLE

The WashU-Merck EST Project

JOURNAL

Unpublished (1995)

COMMENT

Contact: Wilson RK
WashU-Merck EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
High quality sequence stops: 354
Source: IMAGE Consortium, LLNL
This clone is available royalty-free through LLNL; contact the IMAGE Consortium (info@image.llnl.gov) for further information.

FEATURES

source
1..454
Location/Qualifiers
/organism="Homo sapiens"
/clone="151568"

BASE COUNT 144 a 86 c 119 g 102 t 3 others

ORIGIN

Query Match 22.0%; Score 354; DB 16; Length 454;
Best Local Similarity 97.1%; Pred. No. 0.00e+00;
Matches 405; Conservative 0; Mismatches 4; Indels 8; Gaps 6;

Db 39 AGCGGGAACTCTCATGGAGGTAAAGGGAATGAAGAGAAATGAACAGGGCTTGA 98
QY 819 agcgggaaactctcatggaggtaaaaagggatgaagagaaatgaagaggggctga 878
|||

Db 99 GGATGAGAAAGAGAGAGAAAGCCCTGGAAGATGACATAGAGAGCGAAGC-TGCCAGG 157
QY 879 ggatgagaaagagagagaaagccctggaagatgacatagagagagcagctgcgag 938
|||

Db 158 AGATGTGTTTTCCCTTAAGGTGAATGAAGCAGGTGAATTCGGCTGATTCGTGTCGAAAG 217
QY 939 agatgtgttttcccttaagggtgaatgaagcaggtgaatcgctgattctggtgctcaag 998
|||

Db 218 GAAGGCGCTAACCTCCAACTGGAACATAAGCAGATTAATATCTCGCTTGCATGCCAG 277
QY 999 gaagcgctaaactccaaactggaacataag---atttaaatatctcggtgactgcag 1055
|||

Db 278 CTTCAATCATGGGATCTGTGACTGGAAACAGGATAGAGAAGATGATTTTGCATGGATCC 337
QY 1056 cttcaatcatgggactgtgtgactggaaacaggatagagaagatgattttgactggaatcc 1115
|||

Db 338 TGCTGATCGAGATAATGCTATTGCTTCTATATGGCAGTTCGCCCTTGCCAGGTTTAC 397
QY 1116 tgctgatcgagataatgctattgcttcttatatggc-agttccggccttgagaggt-cac 1173
|||

Db 398 AAGAAAGACATTGGCCGCTTTGAACCTCTCTACCTGACCTGGCAACCCCAAGCA 454
 QY 1174 atgaagacattggcc-gattgaactctctctacctgacctg-caaccccaagca 1228

RESULT 8
 LOCUS AA009900 429 bp mRNA EST 01-FEB-1997
 DEFINITION ze83d06.r1 Soares fetal heart NBHH19W Homo sapiens cDNA clone
 ACCESSION AA009900
 NID 91470927
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens

REFERENCE
 AUTHORS Hillier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M., Holman,M., Hultman,M., Kucaba,T., Le.M., Lennon,G., Marra,M., Parsons,J., Rifkin,L., Rohlfing,T., Tan,F., Trevaskis,E., Waterston,R., Williamson,A., Wohlmann,P. and Wilson,R.
 WashU-Merck EST Project
 Unpublished (1995)

TITLE
 JOURNAL
 COMMENT Contact: Wilson RK
 WashU-Merck EST Project
 Washington University School of Medicine
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
 Tel: 314 286 1800
 Fax: 314 286 1810
 Email: est@watson.wustl.edu
 This clone is available royalty-free through LLNL; contact the IMAGE Consortium (info@image.llnl.gov) for further information.
 Insert Length: 1276 Std Error: 0.00
 Seq primer: mob.REGA+ET
 High quality sequence stop: 372.
 Location/Qualifiers
 1. .429
 /organism="Homo sapiens"
 /note="Organ: heart; Vector: pTT73D (Pharmacia) with a modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA was primed with a Not I - oligo(dT) primer [5' TGTTACCAATCAGAGTGGAGCGCGCATCTTTTCTTTTCTTTT 3'], double-stranded cDNA was size selected, ligated to Eco RI adapters (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of a modified pT73 vector (Pharmacia). Library went through one round of normalization to a Cot = 5. Library constructed by M.Fatima Bonaldo. This library was constructed from the same fetus as the fetal lung library, Soares fetal lung NBHL19W."
 /db_xref="taxon:9606"
 /clone="365579"
 /clone_lib="Soares fetal heart NBHH19W"
 /sex="unknown"
 /dev_stage="19 weeks"
 /lab_host="DH10B (ampicillin resistant)"
 <1. .>429

mRNA
 BASE COUNT 135 a 76 c 117 g 96 t 5 others
 ORIGIN

Query Match 20.7%; Score 332; DB 22; Length 429;
 Best Local Similarity 96.8%; Pred. No. 0.00e+00;
 Matches 399; Conservative 0; Mismatches 2; Indels 11; Gaps 11;

Db 1 GAGATAGTTCCAGAGCGGCACTCTCATGGAGGTAAAGGAATGAAGAAATG 60
 QY 805 gagatagtttccagagcgcggaactctctcagggttaaaaggaatgaagagaaatg 864
 Db 61 AAAGAGGGGCTTGAGGATGAGAAAGAGAGAAAGCCCTGAAGATGACATAGAGAG 120
 QY 865 aaagaggggcttgaggatgagaaaagagagaaagccctggaagtacatagaggag 924

Db 121 CGAA-CCTGCGAGGAGATGTTTTCCTTAGGTGAATGAAGCAGGTGAATTCGGCCTG 179
 QY 925 cgaagcctgcgagagagatgttttccctcaagtgaatgaagcagggaattggcctg 984

Db 180 ATTCTCGGTCCAAAGGAAAGCGCTAACTTCCAACTGGAACATAAAGATTTAAATATCTC 239
 QY 985 attct-ggtccaaaggaagcgtaacttccaaactggaacataaagatttaataatctc 1043

Db 240 GGTGACTGCAGCTTCAATCATNGGAGTCTGTGACTGGAAACAGGATAGAGAGATGAT 299
 QY 1044 ggttgactgcagcttcaatcat-gggatctgtgactggaaacagataga-gaagatgat 1101

Db 300 TTGACTGGAATCTGCTGNTGCGAGATAATGCTATTNGGCTTCTATATGCGAGTTCCTCGG 359
 QY 1102 ttgactggaatcctgctgctgacgataatgctatt-ggcttctatggcagttcc-gg 1159

Db 360 CCTTTGGCAGGTCCACAAGAAAGACATTTGGCCCGGATTTGAAACTTCTCTCTA 411
 QY 1160 cctt-ggcaggtc-acatgaagacatt-ggcc-gatt-gaaacttctccta 1206

RESULT 9
 LOCUS R68291 341 bp mRNA EST 01-JUN-1995
 DEFINITION Y106a03.r1 Homo sapiens cDNA clone 138412 5'.
 ACCESSION R68291
 NID 9841808
 KEYWORDS EST.
 SOURCE human clone=138412 library=Soares placenta Nb2HP vector-pT73D (Pharmacia) with a modified polylinker host=DH10B (ampicillin resistant) primer=M13RP1 Rsite1-Not I Rsite2-Eco RI Female placenta obtained at birth (full term). 1st strand cDNA was primed with a Not I - oligo(dT) primer [5' ACTGGAAGATTCGCGCGCAGGAAATTTTTTTTTTTT 3'], double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified pT73 vector. Library went through one round of normalization. Library constructed by Bento Soares and M.Fatima Bonaldo.
 ORGANISM Homo sapiens

REFERENCE
 AUTHORS Hillier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M., Holman,M., Hultman,M., Kucaba,T., Le.M., Lennon,G., Marra,M., Parsons,J., Rifkin,L., Rohlfing,T., Tan,F., Trevaskis,E., Waterston,R., Williamson,A., Wohlmann,P. and Wilson,R.
 The WashU-Merck EST Project
 Unpublished (1995)

TITLE
 JOURNAL
 COMMENT Contact: Wilson RK
 WashU-Merck EST Project
 Washington University School of Medicine
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
 Tel: 314 286 1800
 Fax: 314 286 1810
 Email: est@watson.wustl.edu
 High quality sequence stops: 292
 Source: IMAGE Consortium, LLNL
 This clone is available royalty-free through LLNL; contact the IMAGE Consortium (info@image.llnl.gov) for further information.
 Location/Qualifiers
 1. .341
 /organism="Homo sapiens"
 /clone="138412"

BASE COUNT 95 a 60 c 87 g 98 t 1 others
 ORIGIN

Query Match 20.6%; Score 331; DB 16; Length 341;
 Best Local Similarity 97.9%; Pred. No. 0.00e+00;
 Matches 334; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

Db 1 TACCGCTGGCCGGANGACAAAGTCGGGAAACTTCGAGTGTGTTGTAAGAAACAGTAAACAT 60
 |||||
 QY 1249 taccgctggccggagacaaagtcgggaaacttcgagtggttgtaaaacacagtaacaat 1308
 |||||
 Db 61 GCGCTGTCATGGGAGACACACAGTGCAGATGAAAGTGAAGACAGGGAATTCAG 120
 |||||
 QY 1309 gccctggcaggggagacacagcagtgagatgaaagtggagacaggggaaattcag 1368
 |||||
 Db 121 TTGTATCAAGAACTGATGCTACCAAAAGCATCATTTTGAAGCAGAGAGCGCAAGGCG 180
 |||||
 QY 1369 ttgtatcaagaaactgatctacaaagacatcatttttgaagcagaacgtggcaagg99c 1428
 |||||
 Db 181 AAAACCGCGGAAATCGCAGTGGAGTGGCTGTTGCTGTGTTTACGGTGTATGTCAGATAGC 240
 |||||
 QY 1429 aaaacccggcaatcgcaagtggcgtctgttgcgttttcagcgttatgtccagatagc 1488
 |||||
 Db 241 CTTTTATCTGGGATGACTGATGTTACTATCTATCTTTATATTTGATGTCAGTTC 300
 |||||
 1489 cttttatctgtggannctgaatggtactatcttttatatttgactttgtatgtcagttcc 1548
 |||||
 Db 301 CTGTTTTTTTGTATTCATCATAGGACCTCTGTCAATTT 341
 |||||
 QY 1549 ctggttttttgatgtatgcatacagtaggacctctggtcatttt 1589
 |||||

RESULT 10 T65608 442 bp mRNA EST 07-MAR-1995
 LOCUS YC76b01.r1 Homo sapiens CDNA clone 21599 5'
 ACCESSION T65608
 NID 9574653
 KEYWORDS EST.
 SOURCE human clone-21599 'library=Soares infant brain INIB vector=Lafmid BA host=DH10B (ampicillin resistant) primer=M13RPI Rsite1-Not I Rsite2-Hind III Whole brain from a 73 days post natal female. 1st strand cDNA was primed with a Not I - oligo(dT) primer [5' AACTGGAAGATTCGGCGCGAGAGATTTTTTTTTTTT 3']; double-stranded cDNA was ligated to Hind III adaptors (Pharmacia), digested with Not I and directionally cloned into the Not I and Hind III sites of the Lafmid BA vector. Library went through one round of normalization. Library constructed by Bento Soares and M.Fatima Bonaldo.
 ORGANISM Homo sapiens
 Eucaryotae; Metazoa; Chordata; Vertebrata; Gnathostomata; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 442)
 AUTHORS Hillier, L., Clark, N., Dubuque, T., Elliston, K., Hawkins, M., Holman, M., Hultman, M., Kucaba, T., Le, M., Lennon, G., Marra, M., Parsons, J., Rifkin, L., Rohlfing, T., Soares, M., Tan, F., Trevaskis, E., Waterston, R., Williamson, A., Wohlmann, P. and Wilson, R.
 TITLE The WashU-Merck EST Project
 JOURNAL Unpublished (1995)
 COMMENT GDB: G00-393-946
 Contact: Wilson RK
 WashU-Merck EST Project
 Washington University School of Medicine
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
 Tel: 314 286 1800
 Fax: 314 286 1810
 Email: est@wustl.wustl.edu
 High quality sequence stops: 305
 Source: IMAGE Consortium, LLNL
 This clone is available royalty-free through LLNL; contact the IMAGE Consortium (info@image.llnl.gov) for further information.
 FEATURES
 source Location/Qualifiers
 1..442
 /organism="Homo sapiens"
 /clone="21599"
 BASE COUNT 144 a 81 c 115 g 94 t 8 others
 ORIGIN
 Query Match 19.4%; Score 311; DB 5; Length 442;

Best Local Similarity 95.8%; Pred. No. 0.00e+00;
 Matches 345; Conservative 0; Mismatches 11; Indels 4; Gaps 4;
 Db 6 TATAAGGCAATGGACTTCGGTGTCTGCTATCCCTGAAATCTCTGAGGAAGCCCTC 65
 |||||
 QY 598 tataaaggcaatggacttcggtgtctgtatccctgaaattctctgaaagtcctc 657
 |||||
 Db 66 AGAGACACCTGGGACCATCAAGACAGAAATCAAGAACTGCTGCTCACAAAACAGCATG 125
 |||||
 QY 658 agagacacctggatccatcaaacagaaatcaagaagtctgtctcacaacacagcatg 717
 |||||
 Db 126 AAAAGAAGCGCAAAATTAATAATGTTACCCAGAACCCACAGGNTCTCCCTAAG 185
 |||||
 QY 718 aaaaagaagcgcaaaattaaaattgttaccacagaccaccagagactctacccctaag 777
 |||||
 Db 186 GTGAACTTCGACCCCTTCAACTATGAAGAGATAGTTTCCAGAGCGGGACACTCATGGA 245
 |||||
 QY 778 gtgaacttcgacccctcaactatgaagagatagttccagagggcggaactctcatgga 837
 |||||
 Db 246 GGTAAAAAGGGAATGAAGAGAAATGAAGAGGGGCTTGAGGATGAGAAAGAGAGAG 305
 |||||
 QY 838 ggtaaaaagggaatgaagagaaatgaagagggggcttgaggatgagaaagagaagag 897
 |||||
 Db 306 AAAGCCCTGAAGGATTGACATAGGAGGCGACCTTTTCGAGGGAGATGTTTTTTTCC 365
 |||||
 QY 898 aaagccctgaagaat-gacatag-agg-agcgaagcctgcgagg-agatgtgttttccc 953
 |||||

RESULT 11 AI057064 406 bp mRNA EST 29-SEP-1998
 LOCUS Oz07b04.x1 Soares fetal_liver_spleen_lNFLS_S1 Homo sapiens CDNA
 DEFINITION clone IMAGE:1674607 3', mRNA sequence.
 ACCESSION AI057064
 NID 93330853
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 406)
 AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
 TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index
 JOURNAL Unpublished (1997)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Tel: (301) 496-1550
 Email: Robert.Strausberg@nih.gov
 This clone is available royalty-free through LLNL; contact the IMAGE Consortium (info@image.llnl.gov) for further information.
 Insert Length: 757 Std Error: 0.00
 Seq primer: -40ml3 fwd. ET from Amersham
 High quality sequence stop: 372.
 FEATURES
 Location/Qualifiers
 1..406
 /organism="Homo sapiens"
 /note="Organ: Liver and Spleen; Vector: pT73D (Pharmacia) with a modified polylinker; Site_1: Pac I; Site_2: Eco RI; This is a subtracted version of the original Soares fetal liver spleen INFLS library. 1st strand cDNA was primed with a Pac I - oligo(dT) primer [5' AACTGGAAGAAATTAAGATCTTTTTTTTTTTTTTTT 3'], double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Pac I and cloned into the Pac I and Eco RI sites of the modified pT73 vector. Library went through one round of normalization. Library constructed by Bento Soares and M.Fatima Bonaldo."
 /db_xref="taxon:9606"
 /clone="IMAGE:1674607"
 /clone_lib="Soares_fetal_liver_spleen_lNFLS_S1"
 /sex="male"
 /dev_stage="20 week-post conception fetus"
 /lab_host="DH10B (ampicillin resistant)";

Contact: Wilson RK
WashU-Merck EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
High quality sequence stops: 283
Source: IMAGE Consortium, LLNL
This clone is available royalty-free through LLNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Location/Qualifiers
1..364

FEATURES
source

BASE COUNT 97 a 56 c 95 g 111 t 5 others
ORIGIN
Query Match 15.5%; Score 249; DB 5; Length 364;
Best Local Similarity 96.6%; Pred. No. 0.00e+00;
Matches 281; Conservative 0; Mismatches 5; Indels 5; Gaps 5;
Db 1 GCATGGGAGACACACAGTGAAGATGAAAGTGAAGACAGGGAATTCAGTTGTAT 60
QY 1315 gcatgggagagaccacagtgaggatgaaagtggagagcagggaattcagttgtat 1374
Db 61 CAAGGAACGTGATGCTACCAAGCATCATTTTGAAGCAGACAGTGGCAAGGGCAAAACC 120
QY 1375 caaggaactgctgctaccaaagcatcattttgaagcagacgtggcaagggcaaaacc 1434
Db 121 GCGAAATCGCAGTGGATGGCTGCTGCTTTTCAGGCTTATGTCAGATAGCCTTTTA 180
QY 1435 ggcgaatcgcagtgagtgctgctgcttttcaggtttatgcttgcagatagcctttta 1494
Db 181 TCTGTGGGATGACTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 240
QY 1495 tctgtgg-annnctgaatgtactatttataattgactttgtatgcagttccctggt 1553
Db 241 TTTTGTGATATTGCTATCATTAGGACCTCTGGCATTTTAGGAATTACTAG 291
QY 1554 ttttttg-atattgcatcatagg-acctctgg-cattttaa-aattactag 1600

Search completed: Sat Apr 24 23:15:07 1999
Job time : 2372 secs.